

***Drosophila* Genes Associated with Mushroom Body Enhancers.**

A thesis submitted for the degree of Doctor of Philosophy.

by

Andrew Mounsey

Division of Molecular Genetics

University of Glasgow

September 1997

ProQuest Number: 13818626

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 13818626

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

**GLASGOW UNIVERSITY  
LIBRARY**

11324 (copy 1)

The research reported within this thesis  
is my own work, except where otherwise stated,  
and has not been submitted for any other degree.

A. Mounsey

# CONTENTS

Abbreviations	i
Summary	iv
<b>Chapter 1 Introduction</b>	<b>1</b>
1.1 Introduction	2
1.2 The Mushroom Bodies	3
1.3 Brain Development	4
1.4 Chemical Ablation of the Mushroom Bodies	6
1.5 Biochemical Learning Mutants	7
1.5.1 <i>dunce</i>	7
1.5.2 <i>rutabaga</i>	9
1.5.3 Other Mutations	10
1.6 <i>Drosophila</i> Behaviours	13
1.7 Short-Term Memory: A Biochemical Synthesis	14
1.8 Mushroom Body Circuit Model	16
1.9 P-elements	16
1.9.1 Enhancer-trap elements	18
1.10 Conclusion	21
<b>Chapter 2 Materials and Methods</b>	<b>22</b>
2.1 <i>Drosophila</i>	23
2.1.1 Culture	23
2.1.2 Nucleic Acid Isolation	23
2.1.3 Plasmid-Rescue	24
2.1.4 Sectioning and Staining	25
2.1.5 <i>In situ</i> Hybridisation	25
2.2 Bacterial Protocols	27
2.2.1 Bacterial Strains	27
2.2.2 Plasmids	28
2.2.3 Bacterial Culture Media	28
2.2.4 Antibiotics and Indicators	29
2.2.5 Transformation of <i>E. coli</i>	29
2.2.6 Plasmid DNA Preparation	30

## CONTENTS (cont.)

2.3	Bacteriophage Protocols	31
2.3.1	Bacteriophage Vectors	31
2.3.2	Preparation of Plating Cells	31
2.3.3	Bacteriophage DNA Preparation	31
2.3.4	Screening Phage Libraries	33
2.4	General Molecular Biology Protocols	33
2.4.1	Quantitation of Nucleic Acids	34
2.4.2	Oligonucleotide Synthesis	34
2.4.3	Restriction Enzyme Digests	34
2.4.4	Labelling Nucleic Acids	34
2.4.5	Size Fractionation of Nucleic Acids	36
2.4.6	Blotting and Nucleic Acid Hybridisation	38
2.4.7	Autoradiography	39
<b>Chapter 3</b>	<b>Reverse Northern Analysis of Rescued-Plasmids</b>	<b>40</b>
3.1	Introduction	41
3.2	Isolation of Rescued-Plasmids	41
3.3	Reverse Northern Analysis	42
3.4	Discussion	44
<b>Chapter 4</b>	<b>The 201Y 3' Gene</b>	<b>46</b>
4.1	Introduction	47
4.2	Analysis of the 201Y 3' Gene	47
4.2.1	Reverse Northern Analysis of pB201Y	47
4.2.2	Isolation and Sequence Analysis of cDNA clones	49
4.2.3	Genomic Southern	53
4.2.4	Orientation of the FKBP12 Gene	54
4.2.5	Expression of the FKBP12 Gene	55
4.3	Discussion	57
4.3.1	FK Binding Proteins	57
4.3.2	FKBPs and Calcineurin	59
4.3.3	FKBPs and Ryanodine Receptors	63
4.3.4	FKBPs and Rapamycin	65

## CONTENTS (cont.)

<b>Chapter 5</b>	<b>The 201Y 5' Gene</b>	67
5.1	Introduction	68
5.2	Identification and analysis of the 201Y 5' gene	68
5.2.1	Isolation and Restriction mapping of Genomic clones	68
5.2.2	Reverse Northern analysis of Genomic clones	69
5.2.3	Isolation and Sequence analysis of cDNA clones	70
5.2.4	Genomic Southern	72
5.2.5	Expression of the A421 Gene	73
5.3	Discussion	74
5.3.1	A421 has a Glutamine-Rich Domain	74
5.3.2	A421 has a Novel Zinc-Finger-Like Motif	76
5.3.3	Does the A421 Gene Product Contain a Leucine Zipper?	83
5.3.4	Conclusion	84
<b>Chapter 6</b>	<b>Generation of 201Y Excision Mutants</b>	86
6.1	Introduction	87
6.2	Generation and Analysis of Excision Lines	87
6.3	Discussion	90
<b>Chapter 7</b>	<b>Genes Flanking the 43Y Enhancer-Trap</b>	91
7.1	Introduction	92
7.2	Analysis of the 43Y 3' Gene	92
7.2.1	Restriction Mapping of pB43Y3'	92
7.2.2	Isolation and Sequence Analysis of cDNA Clones	93
7.3	Analysis of the 43Y 5' Gene	95
7.3.1	Isolation and Sequence Analysis of cDNA Clones	95
7.3.2	Restriction Mapping of pB43Y5'	97
7.3.3	Expression of $\alpha$ -Actinin	98
7.4	Discussion	99
7.4.1	<i>Drosophila</i> $\alpha$ -Actinin	99
7.4.2	$\alpha$ -Actinin Neuronal Function	100
7.4.3	$\alpha$ -Actinin and NMDA Receptors	103
7.5	Conclusion	106
	<b>Conclusions</b>	108
	<b>Bibliography</b>	111

## ABBREVIATIONS

### Chemicals:

ATP	adenosine triphosphate
BCIP, X-phosphate	5-bromo-4-chloro-3-indoyl-phosphate
BRdU	5-bromo-2'-deoxyuridine
cAMP	cyclic-3',5'-adenosine monophosphate
cDNA	complementary DNA
DEPC	diethyl pyrocarbonate
DNA	2' deoxyribonucleic acid
dATP	2' deoxyadenosine triphosphate
dCTP	2' deoxycytidine triphosphate
dGTP	2' deoxyguanosine triphosphate
dNTP	2' deoxynucleotide triphosphate
dTTP	2' deoxyuridine triphosphate
DTT	dithiothreitol
EDTA	ethylenediaminetetra-acetic acid (disodium salt)
EtBr	ethidium bromide
hnRNA	heterogeneous nuclear RNA
IPTG	isopropyl-b-D-thiogalactopyranoside
MOPS	3-morpholinopropanesulphonic acid
mRNA	messenger RNA
NaPPi	sodium pyrophosphate
NTB	4-nitrobluetetrazoliumchloride
PEG	polyethylene glycol
RNA	ribonucleic acid
SDS	sodium dodecyl sulphate
TCA	trichloroacetic acid
TE	10mM Tris pH 8.0, 1mM EDTA pH 8.0
Tris	Tris(hydroxymethyl)aminomethane
X-gal	5-bromo-4-chloro-3-indoyl- $\beta$ -D-galactopyranoside

## ABBREVIATIONS (cont.)

### Measurements:

bp	base pair(s)
cm	centimetres
cpm	counts per minute
°C	degrees centigrade
g	grammes
g	force equal to gravitational force
hr	hours
kb	kilobase pairs
kDa	kilodalton
l	litres
mg	milligrammes
mg	microgrammes
min	minutes
ml	millilitres
ml	microlitres
M	molar
mM	millimolar
μM	micromolar
mCi	microcuries
ng	nanogrammes
nm	nanometres
μm	micrometres
nmol	nanomoles
OD	optical density
pH	acidity ( $-\log_{10}[\text{H}^+]$ )
pmol	picomoles
rpm	revolutions per minute
sec	seconds
U	units
UV	ultraviolet light
V	volts
kV	kilovolts
vol	volume
W	watts

## ABBREVIATIONS (cont.)

### Miscellaneous:

DNase I	deoxyribonuclease I
NLS	nuclear localisation signal
ORF	open reading frame
PFU	plaque forming unit
RNase A	ribonuclease A
RF	ribonuclease free

## Summary

The mushroom bodies of *Drosophila melanogaster* are discrete substructures of the fly brain that have been implicated in mediating the processes that underlie learning and memory. Specific chemical ablation of the mushroom bodies greatly reduces flies' performance in associative learning tasks. The two best characterised mutants deficient in learning and memory are *dunce* and *rutabaga*. Both have preferential expression in the mushroom bodies, and have helped elucidate the biochemical mechanisms underpinning learning in *Drosophila*. Identification of other genes with mushroom body preferential, or specific, expression will be necessary to understand other aspects of these biochemical mechanisms.

Enhancer-traps are engineered P-element transposons that carry a reporter gene under the control of a minimal promoter. Genomic insertion of the enhancer-trap in the vicinity of an enhancer allows the expression of the reporter gene, and the identification of tissue specific expression patterns. Genes flanking the enhancer-trap may have a similar tissue specific expression pattern.

Twenty-four fly lines, with enhancer-trap expression in the mushroom bodies, were isolated from a screen of 1400 fly lines containing a single copy of a P[GAL4] enhancer-trap (Yang, 1995, Armstrong, 1995). These twenty-four fly lines were made available to me for further study. Initially, genomic DNA flanking the enhancer-traps was acquired by plasmid-rescue, and analysed by reverse Northern, using probes copied from fly head or body derived mRNA. Many of the rescued-plasmids were shown to contain transcriptional units, some of which clearly had head elevated expression. Two fly lines, 201Y and 43Y, were selected for further detailed study. The genes adjacent to the 201Y and 43Y enhancer-traps were identified by screening cDNA libraries with probes derived from rescued-plasmids or genomic DNA clones, and subjecting isolated cDNA clones to sequencing analysis.

The genes flanking the 201Y enhancer-trap are the previously uncharacterised *Drosophila* homologue of 12kDa FK506 binding proteins (FKBP12), and a new gene (known as A421) with no significant homology to other genes in the DNA sequence databases. Northern analysis revealed that the FKBP12 gene is expressed at all developmental stages, but it was not determined if it has elevated expression in the mushroom bodies. Conceptual translation of the largest A421 ORF gives a protein of 510 aa, that has a glutamine-rich region, a potential leucine-zipper (of novel structure), and a new zinc-finger-like motif of the form WXCX<sub>2-4</sub>CX<sub>3</sub>N<sub>6</sub>CX<sub>2</sub>CX<sub>3-4</sub>R/K. This previously undescribed zinc-finger-like motif was identified in twenty other proteins of diverse functionality, and may play a role in mediating protein-protein interactions. Northern analysis indicates that the A421 gene produces four transcripts, two of which clearly have elevated expression in fly heads. Furthermore, *in situ* hybridisation of DIG

labelled cDNA probes to fly head sections, suggests that the A421 gene is expressed in neurons of the fly brain, but not with elevated or specific expression in the mushroom bodies.

The genes flanking the 43Y enhancer-trap are *ultraspiracle* and  $\alpha$ -actinin. Both of these genes have been previously cloned and characterised in *Drosophila*. The *ultraspiracle* gene product has homology to the retinoid X receptor class of proteins, and interacts with the ecdysone receptor to produce a DNA binding heterodimer that mediates ecdysteroid-induced gene expression (Yao et al., 1993). Immunohistochemical analysis indicates that the *ultraspiracle* gene product is present in all neurons of the fly CNS, with no noticeable concentration in the mushroom bodies.  $\alpha$ -Actinin is an actin cross-linking protein commonly associated with the Z-discs of muscle fibres (Blanchard *et al.*, 1989). However, immunohistochemistry, using an antibody to waterbug  $\alpha$ -actinin, suggests that  $\alpha$ -actinin is preferentially expressed in the mushroom bodies.  $\alpha$ -Actinin is an integral constituent of the neuronal cytoskeleton, and may play a role in the mushroom bodies by mediating neuronal plasticity.  $\alpha$ -Actinin is thought to link NMDA receptors and integrins to the neuronal cytoskeleton and may function in mushroom bodies by regulating the activity of these molecules.

**Chapter 1**  
**Introduction**

## 1.1 Introduction

*Drosophila* is a holometabolous dipteran insect, i.e. it has two wings (the posterior pair of wings forming halteres), six legs emerging in pairs from three thoracic segments, and it undergoes metamorphosis with the fully formed adult or imago emerging from a pupal stage. During the embryonic stage the cells that comprise the body of the larva are formed, and at hatching are fully differentiated and undergo no further division, the rapid growth of the larva being accounted for by increases in cell size. Cells that will later form adult body structures also appear early in embryonic development. They are collected in structures known as imaginal discs and continue to proliferate, but without differentiation, until late larval stages. Early in pupal metamorphosis the larval body structures are degraded and the imaginal discs evert and fuse together. These fused imaginal discs, upon differentiation, form most of the adult epidermal structures such as wings, legs, eyes etc. However, the central nervous system (CNS) is an important exception to this developmental pathway (as are the Malpighian tubules of the excretory system). The brain of the adult *Drosophila* is not formed from an imaginal disc, rather CNS development starts at embryo stage 8, when the first neuroblasts begin proliferating, with new neurons being added continuously until late in pupal development.

The *Drosophila* brain is a complex structure comprising some 200 000 neurons. It is divided into a number of anatomically discrete substructures, whose cell bodies are located in a thin rind on the surface of the brain surrounding a dense neuropil of dendritic and axonal processes. Visual information is initially handled by the optic lobes, structures located in the lateral portions of the brain, before being further processed by the central brain. The optic lobes can themselves be subdivided into the regularly ordered neuropils of lamina, medulla, lobula and lobula plate. The central brain region can be divided into the antennal lobes (used for processing olfactory information), the mushroom bodies (described in detail below) and the central complex. This latter structure can be further subdivided into the ellipsoid body, the fan-shaped

body, the protocerebral bridge, and the noduli. The function of the central complex is not completely clear but it is thought to participate in the coordination and modulation of motor activities (Strauss and Heisenberg, 1993).

## 1.2 The Mushroom Bodies

The mushroom bodies (MBs), or *corpora pedunculata*, are a pair of prominent structures arranged with mirror symmetry in the brains of insects, first identified in honeybees almost one and a half centuries ago (Dujardin, 1850). MB intrinsic neurons, or Kenyon cells (named after F.C. Kenyon, an early neuroanatomist who studied bee MBs), arise from dense clusters of cell bodies above each brain hemisphere. Beneath the cell bodies is a bulbous aggregation of dendrites known as the calyx, and below this the pedunculus, a stalk-like structure of Kenyon cell axons projecting almost to the front of the brain. Here the pedunculus divides into a vertical projecting  $\alpha$ -lobe, and  $\beta/\gamma$ -lobes projecting medially (figure 1.1).

The major inputs impinging upon the calyces of the MBs are from the antennal lobes (AL), which receive olfactory information directly from sensory hairs on the third antennal segment. A bundle of fibres, the antenno-glomerula tract, extends from the AL to the lateral horn of the protocerebral bridge and sends a network of fibres into the MB calyx (Strausfeld, 1976; Stocker *et al.*, 1990). In other insect species the MBs also receive inputs from visual and other sensory systems, and therefore seem in a position to integrate multimodal information. MB outputs extend from the lobes to most major areas of the brain, whereas some fibres provide feedback connections between the calyx, peduncle and lobes, and yet other fibres connect the MBs to each other across the sagittal midplane.

The question arises as to whether the intrinsic neurons of the MB are an isomorphic array of equivalent neurons, or alternatively are these intrinsic neurons genetically or physiologically, and therefore functionally, distinct? A study of enhancer-trap expression within the MBs has revealed some interesting distinctions

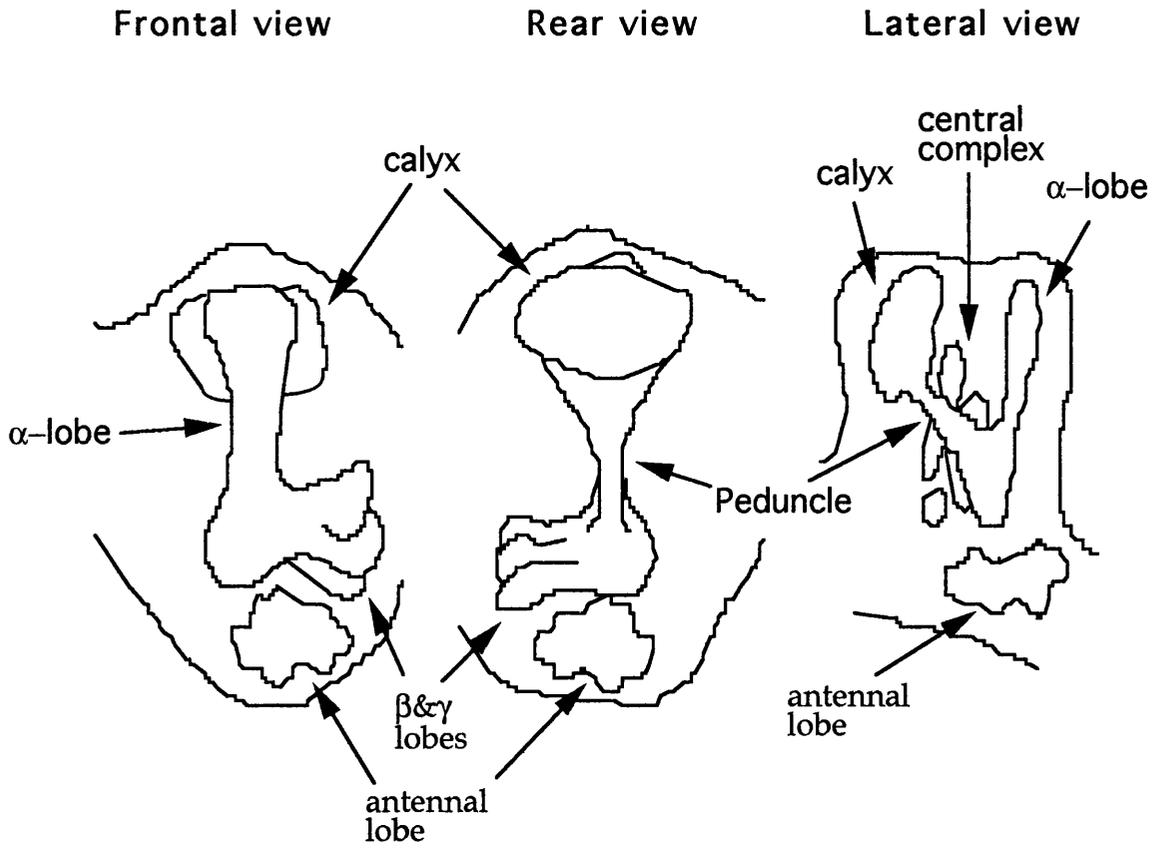


Figure 1.1. Schematic diagram of mushroom body structure. Diagrams outline the major features of one of the symmetrical pair of mushroom bodies. For all views the dorsal side is at the top. For the lateral view, the anterior side is on the right.

(Yang *et al.*, 1995). A detailed examination of eight different fly lines, with P[GAL4] enhancer-trap expression in the MBs, indicate that they are not composed of a genetically homogeneous set of neurons. Each fly line had enhancer-trap expression within the main components of the MBs (i.e. cell bodies, calyx, pedunculus and lobes) but often only comprising a subset of the Kenyon cells involved in these structures. For each line the expression pattern was different, with different lines expressing in different, sometimes complementary, subsets of cells, indicating that at least at the level of gene expression the MBs are composed of distinct sets of neurons.

### **1.3 Brain development**

The postembryonic growth of the *Drosophila* CNS is governed by the activity of a number of neuroblasts, which are considered to be the same as those active in the embryo (Truman, 1990; Prokop and Technau, 1991). Neuroblasts undergo asymmetric division giving rise to a smaller ganglion mother cell (GMC) and another neuroblast. The GMC then divides to form two post-mitotic neurons, which do not undergo any further division. The activity, number and distribution of the postembryonic neuroblasts has been studied in great detail by Ito and Hotta, 1991. By feeding 5-bromo-2'deoxyuridine (BRdU, which is only incorporated by cells undergoing DNA replication) to larvae, or injecting it into the CNS of pupae, they were able to monitor the number and activity of neuroblasts immunohistochemically in whole-mount preparations. They found that immediately after larval hatching there are five active neuroblasts in each hemisphere of the brain. Four are clustered in the dorsal region of the brain above the MB neuropile, and give rise to the intrinsic neurons of the MBs. The fifth neuroblast is at the ventro-lateral side of each hemisphere and gives rise to the antennal lobe structures. No other active neuroblasts appear until eight hours after larval hatching, the number then increasing to approximately 85 per hemisphere some 40 hours later. The number of active neuroblasts then remains constant until 30 hours after puparium formation, when it falls rapidly to leave four per

hemisphere. These four (which are considered to be the same MB neuroblasts observed in the early larval stage) continue proliferating for a further 50 hours until late in pupal development.

The cell cycle period of the MB neuroblasts was determined, for the various developmental stages, by counting the number of surrounding GMCs for different BrdU labelling times. Using these cell cycle data the total number of MB intrinsic neurons, or Kenyon cells, derived from the four persistent neuroblasts was calculated to be 800-1200 per hemisphere. This is in disagreement with the observed number of approximately 2500 Kenyon cell fibres in the pedunculus of adult flies (Balling *et al*, 1987). Although there are 300 Kenyon cells formed during embryonic stages, this is insufficient to account for the discrepancy. A possible explanation is that other neuroblasts observed close to the four MB neuroblasts, during mid larval-early pupal stages, also take part in the formation of the MBs. However, the construction of the *Drosophila* brain, and in particular the MBs, is more complex than a simple addition of neurons throughout development. At the third instar larval stage the number of axons in the peduncle of the MBs is approximately 2100. Within 12 hours of puparium formation, however the number drops by 40% before increasing again to the normal adult level (Technau & Heisenberg, 1982). This degeneration and regrowth affects only the Kenyon cell projections, the cell bodies remaining intact and viable throughout. Two mutants have been described (*mushroom bodies deranged*, *mbd* and *mushroom body defect*, *mud*) that have enlarged MB calyces, but the peduncle and lobes are either completely missing or greatly reduced (Heisenberg, 1980). Interestingly the morphologies of the larval MBs are almost normal (although somewhat enlarged in *mud* mutants), the major defects only appearing in the adult. It would appear that in these mutants the growth of imaginal Kenyon cell projections is somehow misdirected, and the axons remain at the calyces thereby enlarging these structures. Although the MBs undergo a major reorganisation during metamorphosis, a central bundle of approximately 500 larval Kenyon cell fibres appears to remain intact throughout. This central bundle is frequently missing in the *mbd* mutant, and it is

therefore possible that it plays a role in directing the correct growth of the new imaginal fibres. During pupation the peripheral nervous system is completely reorganised, reflecting the demolition and replacement of sense organs and motor systems. The reorganisation of the MBs may be related to these peripheral changes, with the newly connected Kenyon cells enabling the adult fly to process, and respond to, the different environmental cues of adult life.

An examination of the precise role of the individual MB neuroblasts in generating the particular substructures of the MBs has recently revealed further structural divisions within the MBs (Ito *et al.*, 1997). Using clonal analysis, based on a combination of the *flp/FRT* and *GAL4/UAS* systems, they were able to determine that each MB neuroblast gives rise to neurons that contribute to the entire MB structure, i.e. the MBs are fourfold structures composed of clonal sets of neurons derived from each of the four neuroblasts. An examination of 19 enhancer-trap lines with different expression patterns in the MBs, revealed that each line expressed in each of these four clonal sets. This strongly suggests that each clonal unit is a composite of genetically distinct neurons and that each clonal unit also contains most, if not all, of the genetically distinct types of cell comprising the MBs.

#### **1.4 Chemical Ablation of the Mushroom Bodies**

As described in detail in the previous section, for eight hours after emergence of the *Drosophila* larva, only five neuroblasts are active in the *Drosophila* brain. Four of these are MB neuroblasts, and the fifth gives rise to AL neurons. If a larva is fed hydroxyurea (HU, an agent that kills dividing cells) for this eight hour period only, the MBs and some AL neurons will be largely ablated, leaving the rest of the fly brain intact (Prokop and Technau, 1994). Flies that have undergone this treatment have been studied for their performance in an olfactory shock avoidance paradigm (see section 1.6 for a description of this, and other, associative learning paradigms). The HU treated flies had MB calyces some 0.7% of the volume of control flies, with ALs of 68%

normal volume, the rest of the brain being apparently normal (de Belle and Heisenberg, 1994). These flies had almost normal olfactory acuity, shock reactivity and mobility, but were virtual unable to learn an odour-shock association, thereby emphasising the importance of the MBs in associative learning. Further studies using HU ablation have confirmed these results (Jones and Tully, 1995). However, the performance of HU ablated flies was indistinguishable from control flies in both habituation and sensitization of the olfactory jump response. Thus indicating that the MBs are dispensable for the normal functioning of flies in these nonassociative tasks.

## **1.5 Biochemical learning mutants**

It would appear from MB ablation studies (see section 1.4) that the MBs are important in associative learning. A number of single gene mutations have been isolated that also reduce associative learning in flies. Studies of these mutants have helped to confirm the importance of the MBs to associative learning.

### **1.5.1 *dunce***

One of the best understood, most intensely studied, and the first isolated, of the biochemical learning mutants is *dunce*. (Dudai *et al.*, 1976). The first two alleles (*dunce*<sup>1</sup> and *dunce*<sup>2</sup>) generated were EMS induced mutations isolated using odour avoidance learning, and although originally described as learning mutants, careful study revealed almost normal learning but a rapidly decaying memory (Dudai, 1983). In particular, in procedures that couple olfactory cues with food reward, the learning of *dunce* mutants is indistinguishable from that of wild type flies, but again memory decay is rapid (Tempel *et al.*, 1983). Memory through metamorphosis is also deficient in *dunce* mutant flies. Wild type larvae trained to associate an odour with an electric shock in a classical conditioning procedure, will learn the association as larvae and when tested eight days later still possess a memory of it as adults. However, *dunce*

mutant larvae neither learn the association, nor retain a memory of it as adults (Tully *et al.*, 1994). The performance of *dunce* with regard to nonassociative behaviours, such as sensitisation and habituation, is also defective, suggesting a possible conservation of biochemical pathways for associative and nonassociative forms of learning (Duerr and Quinn, 1982). Although *dunce* mutants perform well in tests of olfactory acuity, shock avoidance and motor function, they have at least one other phenotype that highlights the possible pleiotropic effects of biochemical mutants, namely female sterility (Byers *et al.*, 1981).

The *dunce* gene has been cloned, and translation of the cDNA predicts a product with significant homology to known cyclic nucleotide phosphodiesterases (PDE; Chen *et al.*, 1986). A number of studies have shown that all mutant alleles of *dunce* either reduce or abolish Ca<sup>2+</sup>/calmodulin high-affinity cAMP PDE activity (Byers *et al.*, 1981; Davis *et al.*, 1981; Kauvar, 1982; Shotwell 1983). There are at least two other distinct forms of cyclic nucleotide phosphodiesterase in *Drosophila*, a Ca<sup>2+</sup>/calmodulin-sensitive PDE and a cGMP-specific PDE. The *dunce* form of PDE exhibits no specificity towards cGMP even with substrate at a concentration of several hundred micromolar. Expression of the *dunce* gene in yeast gives a form of PDE activity indistinguishable from that missing in *dunce* mutants (Qiu *et al.*, 1991). Interestingly, the *dunce* gene also has weak homology to the precursor of the *Aplysia* egg-laying hormone. Impairment of this function may explain the female sterility, seen in some mutant alleles of *dunce* (Davis and Davidson, 1986).

The *dnc* gene is remarkable in both size and organisation, being the largest characterised *Drosophila* gene. It extends over 148 kilobase pairs, encoding at least 10 RNAs, with the employment of multiple transcription start sites and alternative splicing producing as many as 7 different PDEs (Qui *et al.*, 1991; Qui and Davis, 1993). The precise reasons for this degree of complexity have yet to be worked out, but presumably it is required for the spatial and temporal regulation of multiple PDE isoforms.

In common with other biochemical learning mutants, the possibility arises that the *dunce* phenotype is due to some developmental requirement for the gene product, rather than a physiological one. A number of studies suggest that this is not the case. Feeding adult flies PDE inhibitors produces memory deficits very similar to those found in *dunce* (Folkers and Spatz, 1984; Dudai *et al.*, 1986). More compelling evidence is provided by recent work rescuing the *dunce* memory deficit by expressing PDE in adults under heat-shock control (Dauwalder and Davis, 1995). Interestingly, expression of a rat homologue of *dunce* in adults, was also able to improve the flies memory performance, suggesting functional conservation of *dunce* PDE across species.

Perhaps the most important observation concerning the *dunce* gene product is its expression within the *Drosophila* CNS, and in particular its concentration in the MBs. However, even within the MBs *dunce* PDE is not uniformly distributed. Immunohistochemical experiments reveal it to be concentrated in the  $\gamma$ -lobes and the lateral part of the pedunculus associated with the  $\gamma$ -lobes (Nighorn *et al.*, 1991). Within the medial part of the pedunculus, the calyces and the  $\alpha$ - and  $\beta$ -lobes *dunce* PDE is less concentrated, but still at a higher level than other brain structures. Other studies have revealed that steady-state levels of *dnc* RNAs are elevated in the MB perikarya over other neurons, with the elevated RNAs arising principally from the third transcription start site of the gene (Qui and Davis, 1993).

### **1.5.2 rutabaga**

*rutabaga* (*rut*) is another well characterised learning mutant, originally isolated by an odour avoidance learning screen of EMS induced mutations. As with *dunce* mutants, *rut* mutants show rapid memory loss in a diverse array of associative and nonassociative learning paradigms, from negative reinforcement of olfactory cues, to habituation of the cleaning reflex (Tully and Quinn, 1985; Corfas and Dudai, 1989). Genetic and biochemical studies rapidly associated the *rut* gene with the activity of a

membrane bound,  $\text{Ca}^{2+}$ /calmodulin-sensitive form of adenylyl cyclase (Livingstone *et al.*, 1984; Livingstone 1985; Dudai *et al.*, 1985). Recently, cloning of the *rutabaga* gene has confirmed it as the structural gene for a  $\text{Ca}^{2+}$ /calmodulin-sensitive adenylyl cyclase (Levin *et al.*, 1992). When expressed in a human kidney cell line, the activity of the *rut* gene product was activated by GTP analogues, suggesting a possible regulation by  $G_s$  proteins *in vivo* (Levin *et al.*, 1992). However, absence of this enzyme in *rut* mutants does not reduce basal levels of cAMP, suggesting a regulatory, rather than a maintenance, role for this enzyme (Livingstone *et al.*, 1984). Alternatively the *rut*-encoded enzyme may have a very low basal activity such that enzyme deficiency has subtle effects on steady-state levels, the implication of this being that the activation of the enzyme, by  $\text{Ca}^{2+}$  or signals mediated by G-proteins, is the biochemical property important for learning.

Six new alleles of *rut* have been isolated, from a screen of 5300 fly lines each with a single insertion of a P-element enhancer-trap (carrying a *lacZ* reporter gene). All six were part of a group of 90 fly lines initially isolated for their preferential expression of the enhancer-trap in the mushroom bodies (Han *et al.*, 1992). Each new allele has an insertion in the 5' flanking region of the *rut* gene within 194 bp of the transcription start site. For each allele the severity of the mutation correlates with the proximity of the insertion to the transcription start site, and inversely with the  $\text{Ca}^{2+}$ /calmodulin-responsive adenylyl cyclase activity in fly heads (Han *et al.*, 1992). Furthermore, RNA *in situ* hybridisation and immunohistochemical studies demonstrated that the *rut* gene itself, although expressed throughout the fly brain, has elevated expression in the MBs (Han *et al.*, 1992). In particular, adenylyl cyclase is concentrated in the axonal projections comprising the peduncle and the  $\alpha$ ,  $\beta$ , and  $\gamma$ -lobes, with levels in the calyces being much lower and similar to that found in other brain structures.

### 1.5.3 Other Mutations

A number of other biochemical learning mutations exist, but most have not been studied in such great detail as *dunce* or *rutabaga*. For some the genes have been cloned, and are thus able to help shed some light on the molecular mechanisms underpinning learning and memory.

Several mutations in *DCO*, the structural gene of the catalytic subunit of protein kinase A have been isolated (Lane and Kalderon, 1993; Skoulakis *et al.*, 1993). Heterozygotes of two lethal alleles, *DCO*<sup>B10</sup> and *DCO*<sup>581</sup>, have only 60% wild type protein kinase A activity, whilst a viable complementary cross has only 20% wild type activity. These latter flies have normal olfactory acuity and shock reactivity, and when tested in a negatively reinforced olfactory learning procedure show greatly reduced initial learning, but a stable and longlasting memory (Skoulakis *et al.*, 1993). Similar results were found in transgenic flies expressing a pseudosubstrate peptide inhibitor of protein kinase A (Drain *et al.*, 1991). Most interestingly, the catalytic subunit of protein kinase A, like *dunce* and *rut*, is preferentially expressed in the MBs of the adult fly brain (Skoulakis *et al.*, 1993). The RNA is found principally in the perikarya of MB cells with little or none in the dendritic or axonal processes. Immunohistochemistry experiments revealed that the PKA protein accumulates in the dendrites and axonal projections of MB neurons with low levels found in the remainder of the general brain neuropil.

Dopa decarboxylase is required for the synthesis of the neurotransmitters dopamine and serotonin. Mutations in the gene encoding this enzyme (*Ddc*) are lethal in *Drosophila*, but the existence of temperature-sensitive mutations has allowed the role of this gene in learning and memory to be studied. Mutants that reduce the rate of dopa decarboxylation to approximately 50% reduce the initial performance scores to 40-50% in olfactory avoidance conditioning (Tempel *et al.*, 1984). Whilst other mutants that further reduce enzyme activity have a greater reduction in initial performance scores. Interestingly it has been demonstrated that dopaminergic fibres invade the MB lobes

(Nassel and Elekes, 1992), and this together with the behaviour of *Ddc* mutants, suggest a role for the neurotransmitters dopamine and serotonin in modulating the sensory information presented during training.

The *Su-var(3)6<sup>01</sup>* mutation which effects the structural gene for protein phosphatase 1, and as a homozygote reduces enzyme activity to 30% of wild type, is also deficient in learning and memory (Asztalos *et al.*, 1993). In an olfactory avoidance conditioning paradigm, *Su-var(3)6<sup>01</sup>* mutant flies displayed a large initial decrement in learning and poor memory thereafter. However, although the flies had normal olfactory acuity and shock reactivity, they had reduced motility and flight activity.

*turnip*, a learning mutant isolated in a behavioural screen, has only 20%-30% of normal protein kinase C (PKC) activity (Quinn *et al.*, 1979; Choi *et al.*, 1991). Although not thought to be the structural gene for PKC, *turnip* was considered as a possible regulator of PKC, but genetic analysis suggests it is more complex than this and may not be a monogenic trait (Tully, unpublished observations, cited in Heisenberg, 1989).

An allele of *shaker*, a gene encoding a family of K<sup>+</sup> channels, has been shown to be deficient in learning but almost normal in memory (Cowan and Siegel, 1986). This is interesting in view of the observation that the MBs are rich in *shaker*-encoded K<sup>+</sup> channels (Schwarz *et al.*, 1990).

The learning mutants *radish*, *amnesiac*, *cabbage*, *latheo* and *linotte* have also been isolated in behavioural screens (Aceves-Pina and Quinn, 1979; Folkers *et al.*, 1993; Boynton and Tully, 1992; Dura *et al.*, 1993). Little is known about the molecular nature of *radish*, and *cabbage*, but *amnesiac*, *latheo*, and *linotte* have recently been cloned. In a screen of P-element mutations that rescue the dunce female sterility phenotype, new alleles of *amnesiac* were recovered (Feany and Quinn, 1995). Molecular analysis revealed the P-elements to be inserted in a gene encoding neuropeptides, with some homology to mammalian adenylyl cyclase activating peptide and growth hormone releasing hormone. The *linotte* gene encodes a putative receptor tyrosine kinase, homologous to the human protein RYK (Dura *et al.*, 1995), whilst

cloning of a gene that rescues the lethal phenotype of *latheo*, revealed a transcript with no homology to known sequences (Mihalek *et al.*, 1995). However, these mutants, and *radish* and *amnesiac* in particular, have been useful in dissecting the various phases of memory formation in *Drosophila* (Tully *et al.*, 1990; Tully *et al.*, 1994).

## 1.6 *Drosophila* behaviours

The detailed characteristics of *Drosophila* learning and memory have been studied by analysis of a variety of associative and nonassociative behaviours. Olfactory classical conditioning is a widely used, robust and well tested technique for the evaluation of flies' associative memory. A standard procedure involves placing 100-150 flies in a small cylinder with an electrifiable copper surface, whereupon the flies are simultaneously presented with an odour and a series of small electric shocks. A different control odour is then presented without electric shocks. At various time intervals afterwards, the flies are placed in a T-maze and presented with a choice between the two odours. Memory is quantified as the fraction of flies avoiding the shock-associated odour minus the fraction that does not. Typically, when applied to wild-type Canton-S flies, eighty percent of the flies avoid the shock-associated odour immediately after training, producing a learning index of about 0.6 (Tully and Quinn, 1985). An alternative to this negatively reinforced associative paradigm is the coupling of odour cues with a reward. The flies are similarly presented with two odours sequentially, but in this situation one odour is associated with a sugar solution. Again, after training, the flies are placed in a T-maze and presented with a choice between the two odours, with memory performance being calculated in a similar fashion (Tempel *et al.*, 1983). Interestingly the characteristics of learning in this paradigm are different from those of negative reinforcement, being less robust but more long lived. Similar paradigms associating different colours of light with negative or positive reinforcers, to test visual associative learning, have been developed (Folkers and Spatz, 1981;

Heisenberg, 1989). Associative learning is not confined to adults, larvae can learn to associate olfactory cues and negative reinforcement (Aceves-Pina and Quinn, 1979).

Nonassociative behaviours, such as habituation and sensitisation, have also been studied extensively in the context of learning and memory. Application of weak solutions of sucrose to leg chemoreceptors elicits a reflex extension of the proboscis. Repeated application of the sucrose solution leads to a decrement, or habituation, in proboscis extension events (Duerr and Quinn, 1982). Habituation of the cleaning reflex upon repeated stimulation of the thoracic bristles has also been studied (Corfas and Dudai, 1989). A fly suspended by a fine wire will beat its wings as if in flight. Presentation of a rising horizontal stripe in its visual field will elicit the postural changes required for imminent landing. Repeated presentation of the visual cue leads to habituation of this landing response (Rees and Spatz, 1989). Sensitisation is a nonassociative behaviour that has also been studied in the context of learning and memory. Water applied to the leg chemoreceptors will not normally elicit the proboscis extension reflex. However, if the water is applied after sensitisation by application of a sucrose solution to the proboscis, then the extension reflex is evident (Duerr and Quinn, 1982). Extension of the proboscis within two seconds after the presentation of the water droplet is scored as a positive response.

It is clear that there are a large variety of fly behaviours that can be studied in the context of learning and memory. It is also apparent that a mutant that performs poorly in one behavioural paradigm often performs poorly in others as well, indicating a common physiological link underpinning the different behavioural tasks.

### **1.7 Short-Term Memory: A Biochemical Synthesis.**

Perhaps the most interesting observation arising from the study of biochemical learning mutants, is the fact that three of them, *dunce*, *rutabaga*, and *DCO*, all encode proteins involved in the cAMP second-messenger pathway. This is particularly so, when one bears in mind the model used to explain sensitisation, and associative conditioning of

the gill withdrawal reflex in *Aplysia californica* (Kandel and Schwartz, 1982; Hawkins *et al.*, 1983). Their model of associative conditioning assigns a critical role to the coactivation of adenylyl cyclase. Stimulation of the sensory neuron causes  $Ca^{2+}$  activation of adenylyl cyclase; simultaneously, neurotransmitter release by a modulatory neuron activates adenylyl cyclase via a receptor/G protein complex. This coincident activation of adenylyl cyclase has a more pronounced effect upon enzyme activity than two consecutive activation events (Hawkins *et al.*, 1983). The subsequent prolonged increase in cAMP synthesis induces activation of PKA, by causing dissociation of the catalytic subunits from the inhibitory regulatory subunits, and the consequent cascade of protein phosphorylation alters the physiology of these cells. For example,  $K^+$  channels are closed by phosphorylation, with the result that further action potentials have delayed repolarisation. Thus voltage-gated  $Ca^{2+}$  channels stay open longer, leading to an enhanced response upon firing of the sensory neuron alone (figure 1.2). It is entirely possible that other substrates of PKA, as well as  $K^+$  channels, play a part in this chain of events.

Assuming this scheme is applicable to associative learning in *Drosophila*, it is relative easy to see why the *rutabaga* and *DCO* mutants have defective memory. The lack of adenylyl cyclase would prevent cAMP formation, whilst reduced PKA activity would prevent phosphorylation of the necessary substrates. However, *dunce* mutants have high resting levels of cAMP, and thus require a more subtle explanation of their phenotype. The constantly high levels of cAMP may alter the effectiveness of PKA activation, or lead to desensitisation of the necessary neurotransmitter receptors. The model can also accommodate the observations on *Ddc* and *PPI* mutants. Should dopamine or serotonin be the modulatory neurotransmitters, *Ddc* mutants would be predicted to fail in learning because of the requirement of the enzyme for neurotransmitter biosynthesis. The *PPI* mutants would be defective in the dephosphorylation of essential PKA substrates, with the imbalance producing learning deficits. The transcription factor CREB is activated by PKA phosphorylation, leading to changes in protein synthesis and longer term alterations in cellular physiology. Thus

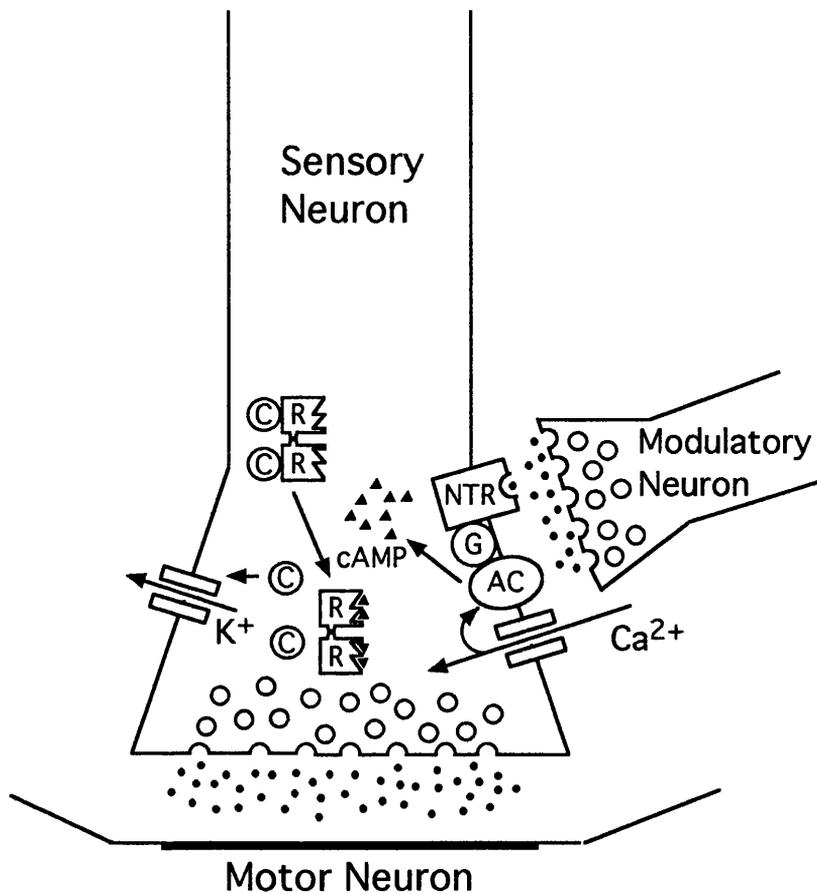


Figure 1.2. Model of associative learning in the Aplysia gill-siphon withdrawal reflex. Stimulation of a modulatory neuron activates adenylyl cyclase (AC) via a G-protein (G) coupled neurotransmitter receptor (NTR). Prior stimulation of the sensory neuron also activates adenylyl cyclase by Ca<sup>2+</sup> causing a prolonged increase in cAMP synthesis. cAMP activation of PKA leads to a cascade of protein phosphorylation, and alteration of the sensory neuron's physiology. Repolarisation is delayed by phosphorylation of K<sup>+</sup> channels, with the result that Ca<sup>2+</sup> channels stay open longer, leading to an enhanced response upon firing of the sensory neuron alone.

mechanisms exist whereby the short term changes centred on the cAMP signalling system can lead to the establishment of long term memories.

It is probable that the cAMP signalling pathway does not act alone and that other mechanisms are at work (Davis, 1993). It would not be surprising if several biochemical pathways are involved in the complex behaviour of associative learning. Perhaps the PKC activity influenced by *tur*, or the unknown products of other mutants, participate in these hypothetical alternative pathways. However, there is no doubt that many molecules expressed in MB cells are likely to be involved in learning and memory. Amongst these we can include transcription factors, proteins involved in synaptic vesicle release and neuronal plasticity, cell surface molecules, as well as those molecules involved in signal transduction cascades. It is clear that a detailed understanding of MB physiology, in normal and learning-deficient flies, is required to elucidate the precise mechanisms underpinning learning and memory in *Drosophila*.

### **1.8 Mushroom Body Circuit Model**

The structure of the *Drosophila* brain, and in particular the position of the MBs within it, suggests the MBs are the site for the integration and possibly storage of information acquired during olfactory learning. *Drosophila* sense odours by receptor hairs on the third antennal segment. Initially this olfactory information is processed by the antennal lobes of the brain before being transferred by projection neurons to the MBs (Davis, 1993). The MBs form part of an olfactory pathway that eventually leads to motor systems (Strausfeld, 1976). It is possible that the MBs act as centres for the integration of information from different sensory modalities, leading to a modulation of the olfactory pathway (figure 1.3). For example neurons carrying the unconditioned stimulus of electric shock, presented during olfactory conditioning, may impinge upon Kenyon cells, altering their biochemistry and physiology and changing their synaptic output. Thus modulatory neurons are able to alter the olfactory information that arrives

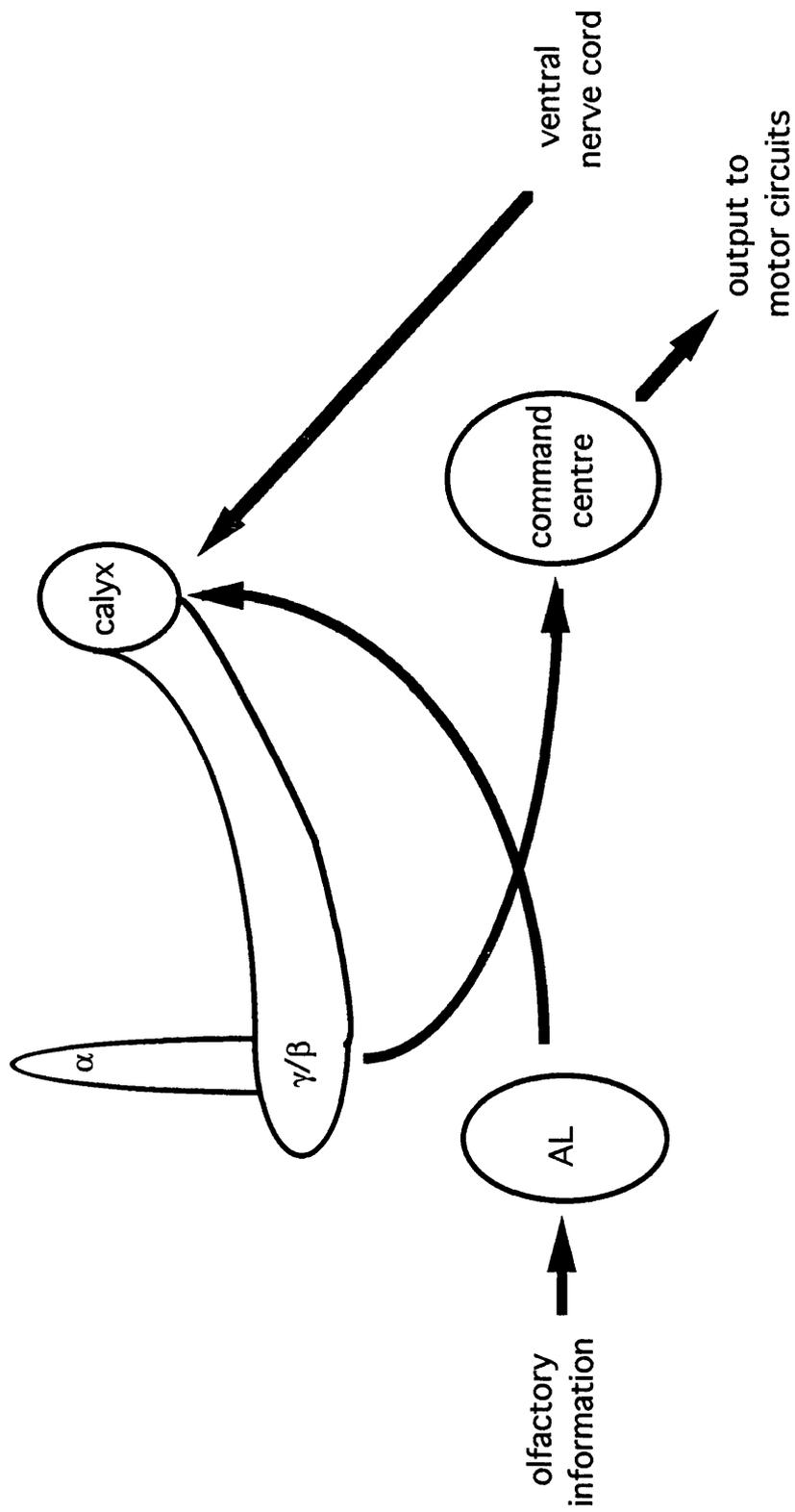


Figure 1.3. Hypothetical model for the modulation of the olfactory pathway by other sensory modalities. Schematic diagram of a sagittal section through the fly brain. Olfactory information is conveyed via the antennal lobe (AL) to the calyx of the MBs. An output is conveyed from the MBs via a command centre to motor circuits. Information from other sensory modalities is carried through the ventral nerve cord to the calyx where it can modulate the olfactory pathway. (Adapted from Davis, 1993).

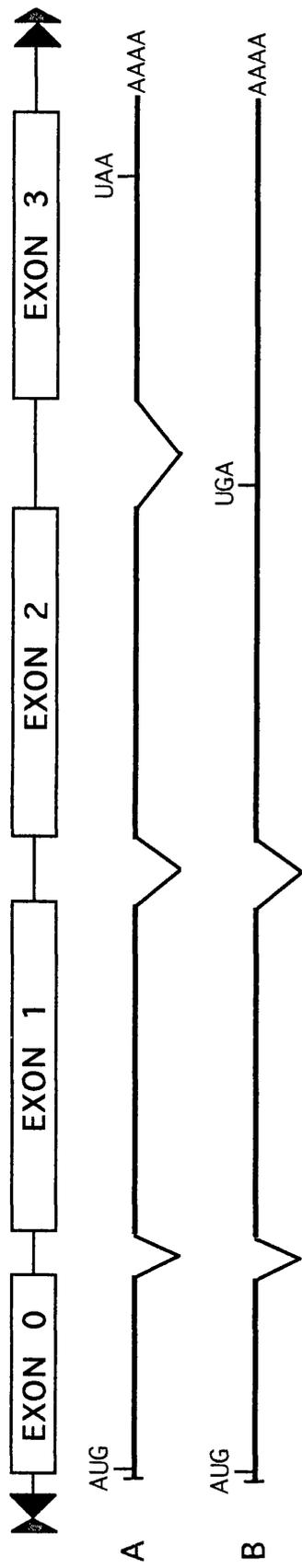
at the dendrites of the Kenyon cells, and carried forward by these cells to their follower neurons and ultimately on to motor systems.

## 1.9 P-elements

P-elements are naturally occurring transposons, that were first described as the causative agents of the syndrome hybrid dysgenesis (reviewed in Engels, 1989). If M strain females (lacking P-elements) are crossed to P strain males (having P-elements), then mobilisation of P-elements occurs in germ-line cells. Phenotypic consequences characteristic of the hybrid dysgenesis syndrome - chromosomal rearrangements, male recombination, sterility, and a high frequency of new mutations - are then evident in later generations. P-elements are 2.9kb in length, with four exons encoding an 87kD transposase, whose activity is restricted to the germline by differential splicing (figure 1.4, Rio, 1991). Internally deleted non-autonomous P-elements, lacking a functional transposase, can still be mobilised in the presence of full-length elements, because they have the requisite cis-acting determinants. This fact has allowed P-elements to be used as tools in mutagenesis strategies.

Improvements on the basic hybrid dysgenesis scheme have accompanied an increased understanding of P-element biology. Fly strains containing specifically engineered P-elements have greatly simplified mutagenesis schemes. The *Birm-2* strain has seventeen non-autonomous P-elements on the second chromosome, and is a useful source of mutagenic 'ammunition' when crossed to a source of transposase (Engels *et al.*, 1987). The P-element derivative P[ry<sup>+</sup> Δ2-3] expresses transposase at high levels, and a strain is available with this element immobilised at a third chromosome site (figure 1.4, Robertson *et al.*, 1988) It is then a simple matter to cross these strains together, create new P-element insertions, and then stabilise these insertions in subsequent generations by "crossing out" the P[ry<sup>+</sup> Δ2-3] containing chromosome. Identification of mutations in a particular gene or genomic region is then dependent on molecular techniques, but is greatly facilitated by the fact that the gene is 'tagged' by the

FULL length P-element



Internally deleted P-element



Δ 2,3

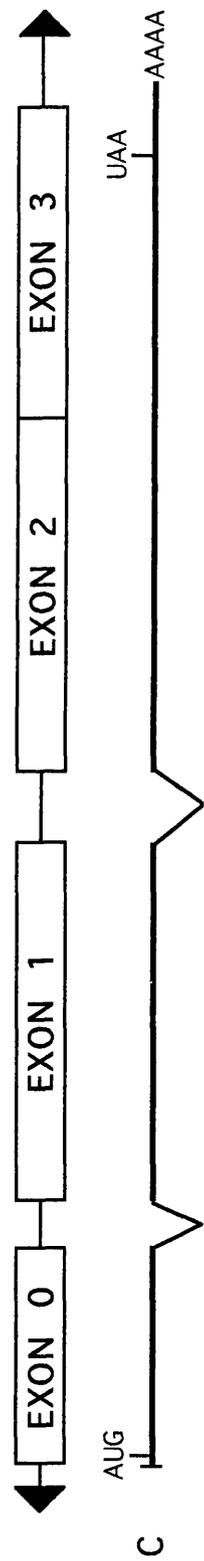


Figure 1.4. Structure of P-elements. Full length P-elements have four open reading frames producing a germline mRNA (A) encoding a 87kDa transposase, or a somatic mRNA (B) encoding a non-functional 66kDa protein. Internally deleted P-elements are present in the 'ammunition' strain *Birm-2*. D 2,3 elements have the third intron removed and produce a mRNA (C) encoding a functional transposase in germline and somatic tissues. (Adapted from Sentry and Kaiser, 1993)

transposon. A strategy based on PCR and sib-selection has been successfully used (Ballinger and Benzer, 1989; Kaiser and Goodwin, 1990), and recently inverse-PCR strategies have been described (Sentry and Kaiser, 1994; Dalby, *et al.*, 1995). However, it may soon be likely that for any gene of interest, a fly line will already exist that has a P-element inserted in or close to it. Large collections of fly lines containing single P-element insertions are already in existence (Cooley *et al.*, 1988; Török, *et al.*, 1992). Screening these collections by PCR, plasmid-rescue or other molecular techniques, should identify any insertions of interest.

### 1.9.1 Enhancer-trap elements

In recent years one of the most productive uses of P-elements has undoubtedly been as enhancer-trap elements, or enhancer-traps for short. Generally, a reporter gene under the control of a minimal promoter, is engineered into an internally deleted non-autonomous P-element (often containing little more than the cis-acting determinants necessary for transposition). Upon insertion at some genomic location, the reporter gene will only be expressed at significant levels if it comes under the control of a nearby endogenous enhancer. The reporter gene expression pattern is assumed to reflect expression of a nearby gene (or genes) under the control of the same enhancer. Early enhancer-traps used the *lacZ* gene of *E. coli*, encoding the enzyme  $\beta$ -galactosidase ( $\beta$ -gal), as a reporter (figure 1.5, O'Kane and Gehring, 1987). Expression of  $\beta$ -gal is visualised by conversion of the chromogenic substrate Xgal, or by using antibodies to the  $\beta$ -galactosidase enzyme. Enhancer-traps also normally carry a phenotypic marker (e.g. *white*<sup>+</sup>), so that flies carrying the enhancer-trap can be easily recognised. Sequences that allow plasmid-rescue of flanking DNA are also included to facilitate rapid identification of nearby genes.

Expression of a transgene in a cell or tissue specific pattern, identified by such enhancer-traps, required tedious cloning and characterisation of flanking DNA to identify sequences that conferred the pattern specificity. Development of a second

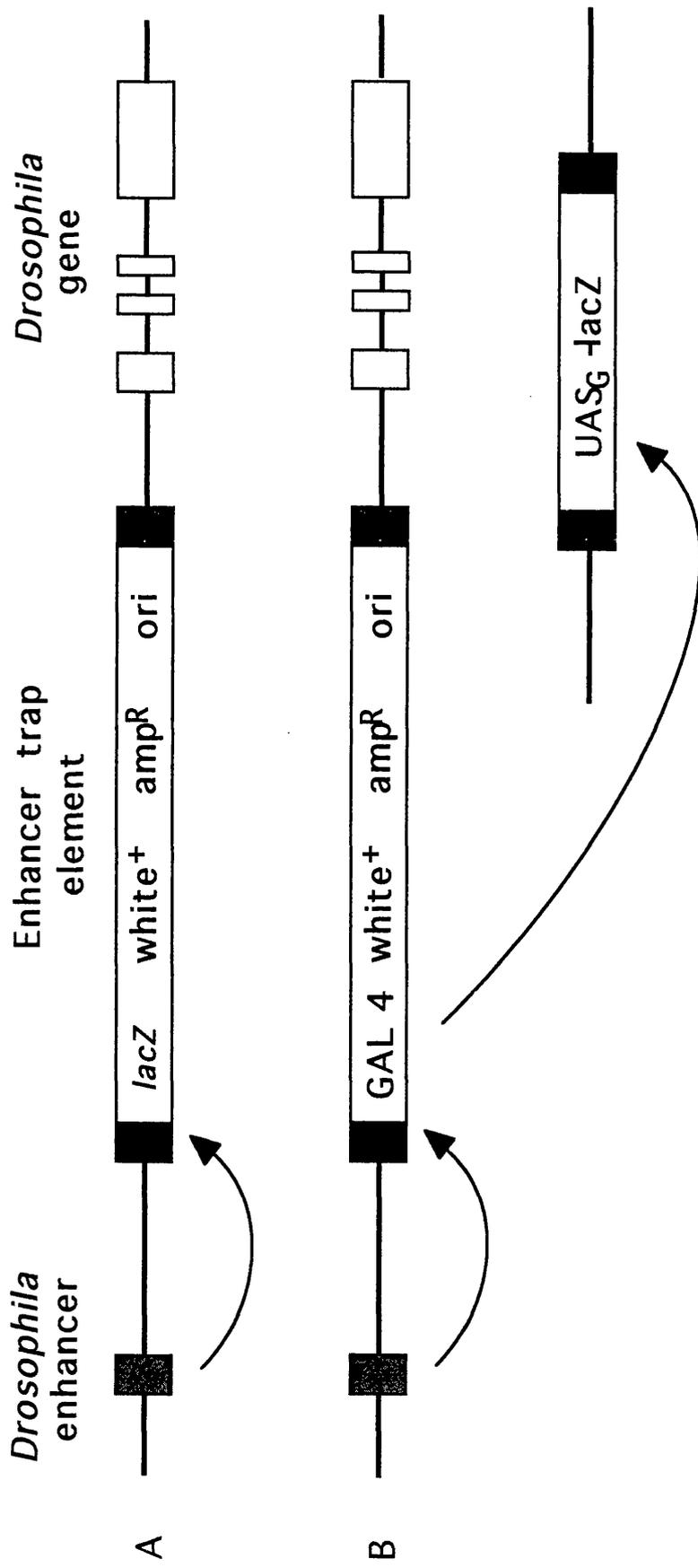


Figure 1.5. Structure of enhancer-traps. A, first generation enhancer-trap, the timing and pattern of expression of the reporter *lacZ* is dependent on a local enhancer. B, GAL4 enhancer-trap. GAL4 expression is visualised by expression of a secondary reporter linked to a GAL4-responsive promoter, UAS<sub>G</sub>. (Adapted from Sentry *et al.*, 1994).

generation of enhancer-traps overcame this problem (Brand and Perrimon, 1993). These new enhancer-traps use a reporter gene that encodes the yeast transcriptional activator GAL4, which has previously been shown to be functional in *Drosophila* (Fischer *et al.*, 1988). Rather than visualise the expression of GAL4 directly, it is used to direct expression of a secondary reporter (e.g. *lacZ*). This secondary reporter is contained within another engineered P-element and its expression is under the control of the GAL4 dependent promoter, UAS<sub>G</sub> (figure 1.5). Once an interesting expression pattern has been identified, any transgene placed under UAS<sub>G</sub> control can be expressed in a similar restricted pattern. Of interest in this respect is the expression of cell ablation agents, such as the polypeptides ricin and diphtheria toxin (Bellen *et al.*, 1992; Moffat *et al.*, 1992). Engineered to be cell-autonomous, and with the development of temperature sensitive versions, these toxins may allow temporally and spatially restricted cell ablation. Of course, many other uses of the “GAL4 system” are possible, including ectopic gene expression, creation of dominant phenotypes, expression of peptide metabolic inhibitors, and modulating gene expression with antisense or ribozyme constructs.

The following example will illustrate the potential uses of this system. Olfactory cues play an important role in *Drosophila* courtship behaviour. In particular, pheromones released by the female are attractive to males (Hall, 1994). The potential role of the MBs in processing this olfactory information in male flies has been investigated by utilising P[GAL4] enhancer-traps (O'Dell *et al.*, 1995). A number of fly lines were isolated with single P[GAL4] enhancer-traps that have expression within subdomains of the MBs. These lines were used to drive expression of the somatic sex-determining gene *transformer* (*tra*), which was placed under UAS<sub>G</sub> control, in the MBs of male flies. Cells expressing *tra* are effectively feminized, and these males, with parts of their MBs consequently feminized, were examined for aberrant courtship behaviour. O'Dell *et al.* found that expression of *tra* in some subdomains of the MBs (but not all) led to indiscriminate courtship behaviour, i.e. other males were courted

almost as vigorously as females. This strongly suggests a role for particular subdomains of the MBs in mate discrimination by males during courtship.

The utility to *Drosophila* neurobiology of the enhancer-trap approach is dependent upon the generation of expression patterns, within the *Drosophila* brain, restricted to particular cell types or anatomical substructures. Only through the generation of such specific expression patterns can the enhancer-trap approach contribute to an understanding of differential gene expression and the dissection of functional sub-domains within the *Drosophila* brain. Are such restricted enhancer-trap expression patterns readily apparent? The results from a large screen of 6000 enhancer-trap lines for expression in the adult brain of *Drosophila*, have recently been reported (Han *et al*, 1996). Of 1082 lines examined in detail 8% showed no staining, whilst 14% had ubiquitous expression in the head, with staining observed in brain, muscles and fat cells as well as retinal and antennal structures. One half of lines had widespread expression within the central nervous system (CNS), classified as; complete CNS staining, or CNS except for the medulla, or staining within the central brain, or within the central brain and the medulla. 16% had specific or preferential expression in limited portions or particular sub-structures of the CNS. The remainder were either unclassified or had expression restricted to peripheral head structures.

Those lines classified as having restricted or preferential expression, were further classified according to which particular structures, within the *Drosophila* brain, showed expression. Of 166 lines examined in detail, 14 lines showed expression in retinal cells or ocelli, 26 in the lamina, and 25 in the medulla. One line showed expression in the antennal lobes, while 22 lines were preferential or specific to the MBs. A further 29 lines had patchy staining in unidentified cells of the protocerebrum. Thus the enhancer-trap approach seems able to generate significant numbers and varieties of restricted expression patterns within the *Drosophila* brain. These expression patterns should prove useful to *Drosophila* neurobiologists in a number of ways. Firstly, they represent valuable markers for cell types or cell populations, within the *Drosophila* brain, for developmental and molecular studies. They can also be used

to identify cell types during *in vitro* culture prior to physiological studies (Wright and Zhong, 1995). These expression patterns have also proved useful as an effective pre-screen for flies with a particular mutant phenotype. For example, a screen of enhancer-trap lines, expressing in the MBs, for deficiencies in learning and memory, has produced numerous alleles of *rut* (Han *et al*, 1992), and one new allele of DCO (Skoulakis *et al*, 1993).

### 1.10 Conclusion

There is clearly good evidence implicating the mushroom bodies as centres of learning and memory in *Drosophila*. Progress has been made in understanding the physiology and biochemistry of mushroom body function, with analysis of the *dunce* and *rutabaga* gene products being particularly revealing, but clearly further work is required. Identification of other genes that have elevated, or specific, mushroom body expression will allow further details of mushroom body function to be elucidated.

A screen of 1400 P[GAL4] enhancer-trap lines had uncovered a number of lines that had preferential expression within the mushroom bodies (Yang, 1995; Armstrong, 1995). Twenty four mushroom body expressing lines were made available to me for molecular analysis, and typical expression patterns for these lines can be seen in figure 1.6. The expression of these lines within the mushroom bodies is not uniform, with different lines staining different subsets of Kenyon cells, and many of the lines also displaying expression in other regions of the fly brain. These enhancer-trap lines represent a potentially important resource in furthering our understanding of mushroom body function. Genes flanking the P[GAL4] elements may also be under the control of a nearby mushroom body enhancer, and will therefore also have preferential expression in the mushroom bodies. The identification and analysis of some of these genes is the purpose of this Thesis.

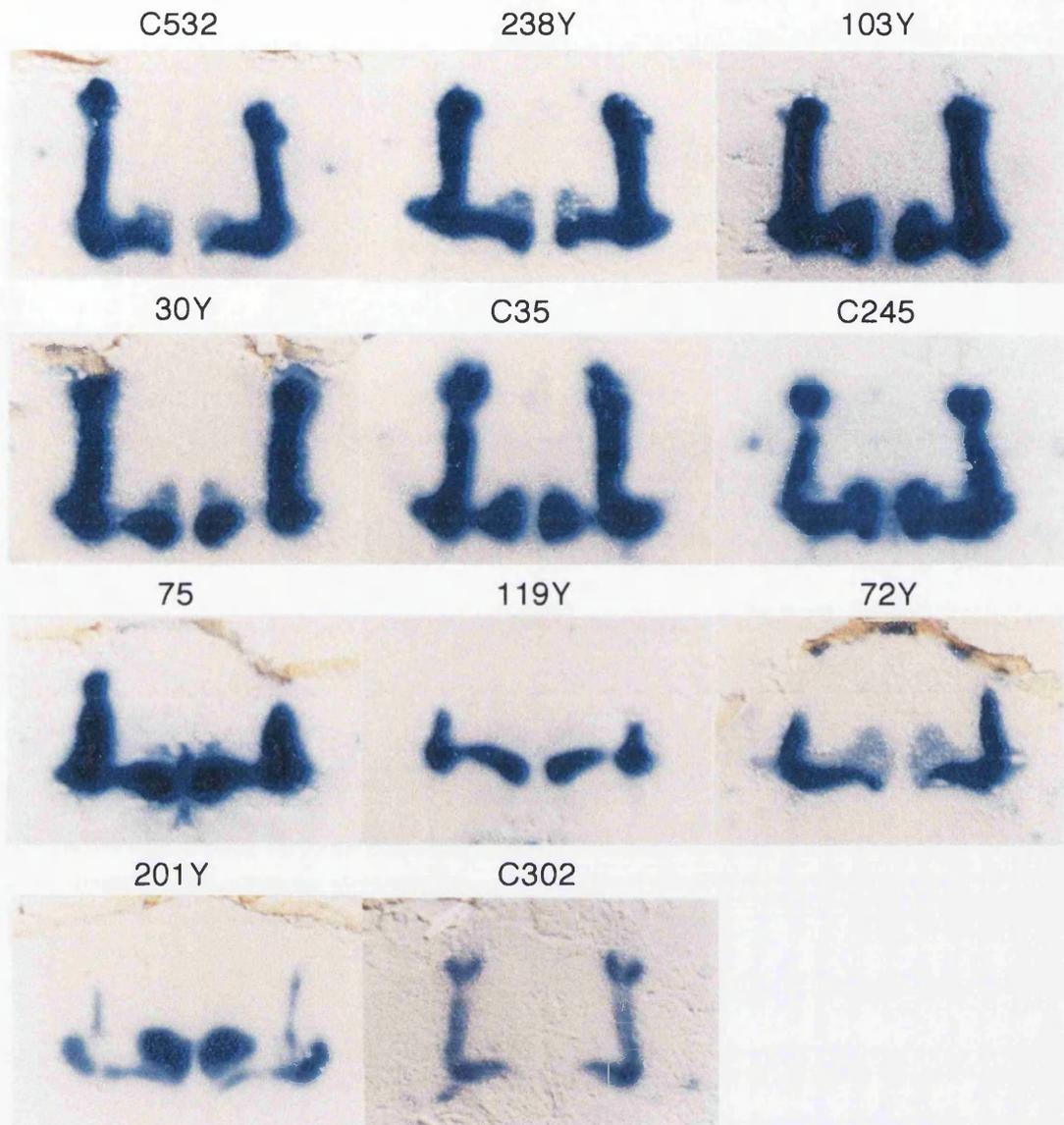


Figure 1.6 Representative expression patterns for some of the fly lines used in this study. Frontal sections of fly heads in the plane of the mushroom body lobes, stained for  $\beta$ -galactosidase activity.

**Chapter 2**  
**Materials and Methods**

## **2.1 DROSOPHILA**

### **2.1.1 Culture**

All flies were routinely cultured at 25°C or 18°C on Glasgow medium (35g dried yeast, 30g glucose, 15g sucrose, 15g maize meal, 10g agar, 10g wheat germ, 10g Soya flour per litre of water). Reagents were dissolved by heating the water to boiling point. Upon cooling to 60°C the antifungal agent Nipagen M (4-hydroxybenzoic acid methylester) was added to a final concentration of 0.1% (w/v).

### **2.1.2 Nucleic acid isolation**

#### **Large-Scale Genomic DNA preparation**

Approximately 1g of flies were added to a liquid Nitrogen cooled mortar and ground to a fine powder in the presence of additional liquid Nitrogen. The powder was transferred into an ice-cold 15ml Wheaton homogeniser containing 9 ml of ice-cold homogenisation buffer (30mM Tris.HCl pH 8.0, 10mM EDTA, 0.1M NaCl, 10mM  $\beta$ -mercaptoethanol) and 500 $\mu$ l of 10% (v/v) Triton X-100. After thorough homogenisation the resulting homogenate was decanted through gauze into a sterile 30ml Corex tube on ice. The homogenate was spun immediately at 4000g for 10 min at 4°C in a cooled rotor. The supernatant was decanted and the nuclear pellet resuspended in 1ml of ice cold homogenisation buffer. 5ml of nuclear lysis buffer (0.1M Tris.HCl pH 8.0, 0.1M EDTA, 0.1M NaCl, 0.5  $\mu$ g/ml Proteinase K.) and 200 $\mu$ l of 30% (v/v) Sarkosyl were added and the solution mixed by swirling prior to incubation for 2 hours at 37°C. The solution was decanted into a preweighed Falcon Tube and 1.25g CsCl per ml of solution was added. The solution was loaded into Polyallomer tubes and filled with 1.25g/ml CsCl. The tubes were sealed and ultracentrifuged at 45K rpm for 16 hours in Ti70 rotor at 25°C. Samples were collected by dripping the solution through

an 18 gauge needle inserted at the bottom of the centrifuge tube. 1.5 ml fractions were collected initially, and then 0.5 ml fractions once the solution became more viscous. The concentration of the DNA samples was crudely estimated using EtBr plates and the fractions with the most concentrated DNA pooled and dialysed at 4°C against several changes of TE. The yields of genomic DNA obtained were generally 100-200µg/gram of starting material.

### **RNA and poly(A)<sup>+</sup> RNA isolation.**

Adult fly heads and bodies were separated and collected using a previously described procedure (Levy and Manning, 1981). One gramme of heads, bodies or whole flies were homogenised in 10mls of denaturing solution (4M guanidinium isothiocyanate, 0.1M Tris-HCl pH8.0, 0.1M β-mercaptoethanol) using a Kinematica Polytron tissue homogeniser for 2-3 mins at full-speed. The homogenate was vortexed after addition of the following: 0.1vol 2M NaOAc, 1 vol. of acidic phenol, and 0.2 vol. chloroform, left on ice for 15-30 min. and centrifuged at 12000g for 10 min. at room temperature. The RNA was precipitated by addition of 1 vol of isopropanol to the aqueous phase, and allowing the mixture to stand at -20°C for 60 min. The precipitate was collected by centrifugation at 12000g for 10 min. at room temperature. The resultant pellet was redissolved in 5ml of denaturing solution, and reprecipitated as described above. The RNA pellet was resuspended in 70% ethanol and placed at -70°C for long term storage, or the RNA was collected by centrifugation, dried and redissolved in 1 ml of H<sub>2</sub>O for immediate use.

#### **2.1.3 Plasmid-Rescue**

Genomic DNA was prepared using a modification of the method of Hamilton *et al.* (1991). DNA was extracted twice with an equal volume of phenol (equilibrated to pH8.0), and once with an equal volume of chloroform, prior to precipitation with an

equal volume of isopropanol. 1µg of genomic DNA was digested for 3hrs at 37°C with 10U of an appropriate restriction enzyme, followed by heat denaturation at 70°C for 15 mins. DNA was ethanol precipitated and redissolved in 100µl of T.E. prior to ligation for 16 hrs at RT with 1U of T4 DNA ligase (Promega) in 1 X ligation buffer and a final volume of 200µl. Following ethanol precipitation, ligated DNA was washed 3 X with 1ml of 70% ethanol, and redissolved in 10µl of distilled water. 5µl was electrotransformed into 40µl of electrocompetent *E. coli* (strain MC1061) and selected on LB agarose plates containing 100µg/ml ampicillin.

#### **2.1.4 Sectioning and Staining for β-Galactosidase Activity**

Fly bodies were in soaked in OCT 4583 medium for 3-5 mins prior to freezing and mounting on a cryotome (Anglia Scientific, Cryotome 620). Sections (8-15 µm) were cut at -19°C and collected onto gelatinised slides prior to air drying at room temperature. Sections were fixed in 1% glutaraldehyde in PBS for 10-15 min at 4°C, and washed three times for 10 min in PBS. Slides were placed in a humid box and 200µl staining solution (0.2% X-gal, 10mM NaPO<sub>4</sub> pH7.2, 0.15M NaCl, 1mM MgCl<sub>2</sub>, 3mM K<sub>4</sub>(FeCN<sub>6</sub>), 3mM K<sub>3</sub>(FeCN<sub>6</sub>), 0.3% (v/v) Triton X-100) added to each with a coverslip placed on top to spread the solution over the specimen. Slides were left at 37°C for 2 hrs or until staining was obvious.

#### **2.1.5 *In situ* hybridisation**

##### **Tissue Sections**

Wild type flies were mounted side by side in a fly collar, soaked in O.C.T. compound (Lab-Tek division) for 30 min., frozen at -18°C, cut as 10 µm sections in a cryotome (Anglia Scientific, Cryotome 620), placed on gelatinised slides and postfixed in freshly made PLP fixative (4% (w/v) paraformaldehyde, 0.01M sodium meta-periodate,

0.075M lysine, 0.044M NaCl, 0.037M phosphate buffer, pH7.4), they were treated with 0.2N HCl for 20 min., washed in 2 X SSPE (0.36M NaCl, 0.02M NaH<sub>2</sub>PO<sub>4</sub>, 2mM EDTA, pH 7.8) for 30 min. at 65°C, treated with 350µg/ml of auto-digested pronase in 50mM Tris-HCl and 5mM EDTA, pH 8.0, incubated in 2mg/ml glycine in PBS, fixed in 4% paraformaldehyde in PBS for 20 min., and acetylated with 0.25% (v/v) acetic anhydride and 0.1M triethanolamine in PBS for 8 min. The sections were incubated with 150-200µl of prehybridisation solution (5 X SSPE, 50% (v/v) formamide, 5% (w/v) dextran sulphate, 1 X Denhardt's solution, 250µg/ml of yeast tRNA, 500µg/ml of sheared Salmon sperm DNA) at 42°C for 2-3hr. They were then incubated overnight with single stranded DNA probes at 42°C. The sections were washed extensively, three times (15 min each) in 2 X SSPE, once in 1 X SSPE, and once in 0.5 X SSPE at room temperature. A final wash was performed in low salt buffer (2mM NaPPi, 1mM NaPO<sub>4</sub>, 1mM EDTA, pH7.2) at 42°C. For immunological detection of the hybridised probes, the sections were incubated with 200µl of 2% (v/v) sheep serum (Gift from S.A.P.U. Law Hospital, Carlisle, Scotland) in buffer I (100mM Tris-HCl, pH 7.5, 150mM NaCl) with 0.3% (v/v) triton X-100 for 60 min. The sections were incubated with 150-200µl of anti-digoxigenin antibody conjugated with alkaline phosphatase (Boehringer Mannheim) diluted at 1:500 in buffer I for 2-3 hrs. After extensive washes in 100mM Tris-HCl, pH9.5, 100mM NaCl, 50mM MgCl<sub>2</sub>, the sections were placed with 200µl of diluted chromogenic substrate solutions (NTB and BCIP, X-phosphate) following the manufacturer's instructions (Boehringer Mannheim), incubated in the dark at room temperature for 2-6 hrs. The reaction was stopped by washing in 10mM Tris-HCl, 1mM EDTA, pH8.0 for 10 min., washed in H<sub>2</sub>O, dehydrated, and mounted with glycerol gelatin (Sigma).

### **Whole-mount brains**

Intact adult brains were removed from fly heads under PBS, fixed in 4% paraformaldehyde in PBS for 30 min, and washed twice for 1 hour in PAT (1 X PBS,

1% bovine serum albumin, 1% Triton X-100). Brains were incubated overnight with primary antibody diluted in 3% normal goat serum in PAT. Brains were washed three times in PAT for 1 hour, incubated overnight with fluorescein-labelled secondary antibody diluted 1:250 in PAT, washed twice for 1 hour in PAT, and once for 5 minutes with PBS. All of the above procedures were carried out at room temperature. Stained brains were mounted in VectaShield (Vector), and visualised on a Molecular Dynamics Multiprobe laser scanning confocal microscope using excitation (480nm) and emission (530+/-15nm) barrier filters appropriate to the fluorescein label used.

## 2.2 BACTERIAL PROTOCOLS

### 2.2.1 Bacterial Strains

The bacterial strains used in this study were XL1-Blue, NM621, and MC1061. Their genotypes are given below.

<i>Strain</i>	<i>Genotype</i>	<i>Reference</i>
XL1-Blue	<i>recA1, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lac, [F' proAB, lacI<sup>q</sup>ZDM15, Tn10(ter<sup>r</sup>)]</i>	Bullock <i>et al</i> , (1987)
NM621	<i>hsdR, mcrA, mcrB, supE44, recD 1009</i>	Whittaker <i>et al</i> , (1988)
MC1061	<i>F<sup>-</sup>araD139, D(ara leu)7697, D(lac) c74, galE15 galK16, hsdR2, strA, mcrA, mcrB1</i>	

### 2.2.2 Plasmids

Plasmids used in this study, other than those whose construction is described within this thesis, are listed below.

Plasmid	Description	Source/Reference
pBluescript®IISK+/-		Mead <i>et al.</i> , (1985) Stratagene USA
pBluescript®IISK+/-		Mead <i>et al.</i> , (1985) Stratagene USA
pST41	<i>EcoRI-HindIII</i> fragment of <i>nina E</i> , major opsin gene of <i>D.melanogaster</i> in pBluescript®IISK-	Gift from S. Tomlinson.

### 2.2.3 Bacterial Culture Media

L-Broth: 10g Bacto-tryptone (Difco), 5g yeast extract (Difco), 10g NaCl, per litre of water and adjusted to pH 7.0 with NaOH.

L-Agar: As above with the addition of Bacto-agar (Difco) to 1.5% (w/v).

0.7% (w/v) Top Agarose: 0.7g agarose added to 100ml of L-broth, containing 10mM MgSO<sub>4</sub>

All culture media were sterilised by autoclaving at 120°C for 15min.

#### **2.2.4 Antibiotics and indicators**

When necessary either ampicillin, at a final concentration of 50 $\mu$ g/ml (50mg/ml stock solution in sterile distilled water), or tetracycline, at a final concentration of 12.5 $\mu$ g/ml (12.5mg/ml stock solution in absolute ethanol) were added to broth or agar. 5-bromo-4-chloro-3-indoyl- $\beta$ -D-galactopyranoside (X-gal) and iso-propyl- $\beta$ -D-thiogalactopyranoside (IPTG) were added to molten agar (50°C) in order to detect recombinant clones. X-gal was dissolved in dimethylformamide, and IPTG in sterile distilled water. Both were stored at -20°C as 20mg/ml, and used at a final concentration of 20 $\mu$ g/ml.

#### **2.2.5 Transformation of *E. coli***

##### **Preparation of competent cells**

A 0.5ml inoculum from an overnight culture of XL1-Blue was added to 50ml of L-broth, and grown with vigorous shaking for approximately 2hrs or until the cells had reached an OD<sub>600</sub> = 0.4. The cells were pelleted by centrifugation at 4000g for 10min. at 4°C, the supernatant discarded, and the bacterial pellet resuspended in 20ml of ice-cold 50mM CaCl<sub>2</sub>. The cells were repelleted as above, followed by resuspension in 2ml of ice-cold 50mM CaCl<sub>2</sub> for immediate use or after 24hrs storage at 4°C.

##### **Transformation**

10 $\mu$ l of a solution of DNA (50-150ng) was added to 200 $\mu$ l of competent cells and left on ice for 45 min. The mixture was heat-shocked at 42°C for 2 min, then placed on ice for 1min. before the addition of 1ml L-broth. The cells were allowed to recover at 37°C for 30min before plating on L-agar plates containing appropriate antibiotics and indicators. Plates were incubated overnight at 37°C to select for transformants.

## **Preparation of electrocompetent cells**

One litre of LB medium was inoculated with a 1/100 volume of a fresh overnight culture of *E. coli* (MC1061). Cells were grown at 37°C with vigorous shaking to an OD<sub>600</sub> of 0.5. Cells were chilled on ice for 15 mins, and centrifuged at 4000g for 15mins at 4°C. The pellet was resuspended in 1 litre of ice-cold water and centrifuged as above. Cells were resuspended in 20ml of ice-cold 10% glycerol and again centrifuged as above. Cells were resuspended in a final volume of 2ml of ice-cold 10% glycerol, and used immediately or stored at -70°C.

## **Electrotransformation**

Electrocompetent *E. coli* were thawed at room temperature and 40µl of the cell suspension added to 5µl of DNA solution on ice, and left to sit on ice for a further one minute. The mixture was transferred to an ice-cold 0.1cm electroporation cuvette, placed within the *E. coli* Pulser® (BioRad) and given one pulse of 1.8 kV. 1 ml of SOC medium (2% Bacto tryptone, 0.5% Bacto yeast agar, 10mM NaCl, 10mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub>, 20 mM glucose) was added immediately to the cuvette and the cells gently resuspended. The cells were transferred to a fresh tube and incubated at 37°C for 1 hour, followed by plating on LB agar plates with the appropriate antibiotics.

### **2.2.6 Plasmid DNA preparation**

The alkaline lysis method of Birnboim and Doly, (1979), as outlined in Sambrook *et al.* (1989) was used for the large-scale preparation of plasmid. Small-scale plasmid preparation was performed using the Magic™ DNA purification system (Promega), following the manufacturers recommended protocol without modification.

## 2.3 BACTERIOPHAGE PROTOCOLS

### 2.3.1 Bacteriophage vectors

The lambda phage vectors used in this study were NM1149 and  $\lambda$ Gem-11. Their genotypes are given below.

NM1149:  $\lambda b538$   $srl\lambda 3^\circ$   $imm434$   $srl\lambda 4^\circ$   $shndIII\lambda 6^\circ$   $srl\lambda 5^\circ$ . Murray, (1983)

$\lambda$ Gem-11:  $shb1$   $\lambda 1^\circ$   $b189$  KH54  $chiC$   $srl\lambda 4^\circ$   $nin5$   $srl\lambda 5^\circ$ . Promega.

### 2.3.2 Preparation of plating cells

To 100ml of LB media plus 0.2% maltose was added 100 $\mu$ l of an overnight culture of bacterial strain NM621. The culture was grown at 37°C with vigorous shaking until the OD<sub>600</sub> was approximately 0.5 and then placed on ice. The cells were centrifuged at 4000g for 5 min at 4°C, and the pellet resuspended in ice-cold 10mM MgSO<sub>4</sub> to a final OD<sub>600</sub> of 1.0. The cells were used immediately or stored at 4°C for up to two weeks.

### 2.3.3 Bacteriophage DNA preparation

#### Moderate scale Bacteriophage DNA preparation

To 500 $\mu$ l of plating cells was added 10<sup>6</sup> PFUs, and the mixture incubated at 37°C for 30 min to allow the phage to preadsorb onto the bacteria. The mixture was added to 37ml of LB medium and grown at 37°C with vigorous shaking for 12-15 hrs or until bacterial lysis was apparent. 100 $\mu$ l of chloroform was mixed thoroughly with the lysed bacteria, followed by the addition of 370 $\mu$ l of nuclease solution (5 $\mu$ g/ml DNase I, 5 $\mu$ g/ml RNase A, 50% glycerol, 30mM sodium acetate pH 6.8) and incubation at 37°C for 30 min. Addition of 2.1g of NaCl was followed by centrifugation at 4000g for 20

mins at 4°C. The supernatant was transferred to new tubes containing 3.7g PEG 6000, and the PEG dissolved by gentle rotation. The tubes were placed on ice for 60 mins, followed by centrifugation at 4000g for 20 mins at 4°C. The precipitated phage were resuspended in 500µl phage buffer (10mM TrisHCl pH 7.5, 10mM MgSO<sub>4</sub>) by gentle rotation. The phage were lysed by addition of 20µl 0.5M EDTA, 5µl of 20% SDS, and 2.5µl of 10mg/ml proteinase K, and incubation at 65°C for 30 mins. The mixture was extracted with phenol (equilibrated to pH 8.0), and then with chloroform. The DNA was precipitated by addition of 700µl of isopropanol, and storage at -20°C for 30 mins. After centrifugation at 12000g for 10 mins at 4°C, the DNA pellet was washed with 1ml of 70% ethanol, dried, and redissolved in 300µl of TE. Yields were typically 50-100µg.

#### **Small-scale Bacteriophage DNA preparation (Plate lysate method).**

Approximately 10<sup>5</sup> pfu were added to 150µl of NM621 plating cells and incubated at RT for 20 min. 3ml of 0.7% (w/v) top agarose were added and the mixture poured onto an LB-agar plate. Plates were incubated at 37°C for 8 hours, or until phage plaques were confluent. Phage were eluted by addition of 5ml phage buffer (10mM Tris.Cl pH 7.5, 10mM MgSO<sub>4</sub>) with gentle agitation for 2 hours. The phage buffer was decanted into an oakridge tube, centrifuged at 15000g for 5 min., and 10µl of RNase/DNase solution added to the supernatant followed by 30 min incubation at RT. 4ml of a 20% (w/v) PEG, 2.5M NaCl solution were added and the solution left on ice for 60 min. This was centrifuged at 15000g for 10 min to pellet the phage particles, the supernatant removed, and the phage pellet resuspended in 500µl of phage buffer. 5µl of 10% (w/v) SDS and 5µl of 0.5M EDTA were added and the solution incubated at 70°C for 15 min. The solution was extracted with an equal volume of phenol, followed by phenol/chloroform and chloroform extractions. The phage DNA was precipitated at room temperature for 30 min by the addition of 350µl of isopropanol. The solution

was centrifuged for 30 min and the resulting pellet washed in 70% ethanol before being resuspended in 50µl TE. The yield obtained was typically 5-10µg of DNA.

#### **2.3.4 Screening Lambda phage cDNA and genomic Libraries.**

Screening of lambda cDNA and genomic libraries was essentially as described in (Sambrook *et al.*, 1989). Briefly, five to ten thousand recombinant phage were plated onto 10cm by 10cm L-agar plates, using  $10^8$  host plating cells and 0.7% (w/v) L-agarose containing 10mM MgSO<sub>4</sub>. The plates were incubated at 37°C for 8hr or until phage plaques were just visible. Replica nylon membranes (Hybond-N) were lifted from each plate and placed plaque side up on 3MM paper soaked in denaturing solution (1.5M NaCl, 0.5M NaOH) for 3 min. The membranes were neutralised on 3MM paper soaked in neutralising solution (1.5M NaCl, 0.5M Tris.HCl pH 7.2, 1mM EDTA) twice for 5 min and washed in 5xSSC for a minimum of 15 min. Then subject to UV crosslinking and probed as described elsewhere.

### **2.4 GENERAL MOLECULAR BIOLOGY PROTOCOLS**

#### **2.4.1 Quantification of nucleic acids**

DNA and RNA was quantified spectrometrically by measuring the absorbance of the sample at a wavelength of 260nm, and assuming for a 1cm pathlength that OD<sub>260</sub> = 1 corresponds to 50µg/ml for DNA and 40µg/ml for RNA. For small volumes of DNA the concentration was estimated by spotting 1µl onto a 1% (w/v) agarose gel containing 0.5 µg/ml EtBr, and comparing the intensity of fluorescence of the sample, upon UV illumination (254nm), to that of known standards

#### **2.4.2 Oligonucleotide synthesis**

Oligonucleotides were synthesised by the solid state method on an Applied Biosystems Inc. PCR-MATE 391 DNA Synthesiser, employing phosphoramidite chemistry. After ammonium hydroxide cleavage and deprotection, oligonucleotides were evaporated to dryness under vacuum and redissolved in water. Typically, sequencing primers were 18 nucleotides in length.

#### **2.4.3 Restriction enzyme digests**

DNA was digested in the appropriate BRL REact buffer for the enzyme, using at least 2U enzyme per  $\mu\text{g}$  DNA. Incubations were carried out at  $37^\circ\text{C}$ , for either 1hr for plasmid and phage DNA, or 3hr for genomic DNA.

#### **2.4.4 Labelling nucleic acids**

##### **First Strand cDNA Probe.**

To  $1\mu\text{g}$  of polyA<sup>+</sup> RNA was added  $70\mu\text{Ci}$  of  $600\text{ Ci/mmol}$  of  $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ ,  $4\mu\text{l}$  of 5x first strand cDNA synthesis buffer (250mM Tris.HCl pH 8.3, 700mM KCl, 50mM  $\text{MgCl}_2$ , 50mM DTT),  $1\mu\text{l}$  80mM NaPPi, 30U RNAGUARD,  $2\mu\text{l}$  Oligo (dT)<sub>12-18</sub>, 20U AMV reverse transcriptase and made up to a final volume of  $20\mu\text{l}$  using RF water. The reaction was incubated at  $42^\circ\text{C}$  for 30 min, followed by the addition of  $1\mu\text{l}$  of 10mM dCTP and the incubation continued for a further 30 min. Hydrolysis of the RNA was achieved by the addition of 1 volume of 0.6M NaOH, 20mM EDTA followed by incubation at  $65^\circ\text{C}$  for 30 min. Unincorporated nucleotides were separated using Sephadex G-50 columns (Sambrook *et al.*, 1989). Incorporation was assessed in a scintillation counter using Cherenkov counting on a small aliquot of the probe. Specific activities of approximately  $1 \times 10^8$  cpm/ $\mu\text{g}$  RNA were normally obtained.

## Random Prime Labelling of DNA

Linearised plasmids or gel-purified DNA fragments were radioactively labelled by random priming (Sambrook *et al.*, 1989). 50-100ng of DNA in (10µl of distilled water) was denatured by heating at 100°C for 5min followed by quenching on ice. The DNA was added to 6µl of 4 X random priming buffer (250mM Tris-HCl, 25mM MgCl<sub>2</sub>, 100mM dNTPs, 1M HEPES pH 6.6, 27 units (A<sub>260</sub>)/µl random hexanucleotides), 50µCi of 3000 Ci/mmole of [ $\alpha$ -<sup>32</sup>dCTP] and 5U of Klenow enzyme (Promega). The reaction was performed in a final volume of 24µl for at least 12hr at RT or 3hr at 37°C. Labelled DNA was separated from unincorporated nucleotides by Sephadex G50 (Pharmacia) chromatography, in columns prepared from disposable 1ml syringes (Sambrook *et al.*, 1989). Incorporation of radioactive precursor was calculated by precipitation of the acid insoluble fraction in TCA. Typically probes had a specific activity of 10<sup>8</sup>-10<sup>9</sup>cpm/µg.

## DIG labelled single-stranded DNA probes

Single-stranded digoxigenin labelled DNA probes were made using a method employing thermal cycling (Patel and Goodman, 1992). The reaction contained 250ng linearised plasmid in 0.5mM KCl, 10mM Tris-HCl, pH8.3, 1.5mM MgCl<sub>2</sub>, 0.01% (w/v) Triton X-100<sup>®</sup>, 0.2mM dATP, 0.2mM dCTP, 0.2mM dGTP, 0.13mM dTTP, 0.07mM Digoxigenin-11-dUTP (Boehringer Mannheim), 150ng of either T7 or T3 primer, 1.25U of *Taq* polymerase (Promega) and was incubated for 30 cycles in a Hybaid Thermal reactor (Hybaid) at 95°C for 45 sec, 55°C for 30 sec and 72°C for 1min. The reaction product was precipitated by adding 10µg of glycogen, 0.1 vol of 3M NaOAc, pH5.0, 3vol of absolute ethanol and storage at -20°C for 1-2 hr. The DNA was pelleted (12,000g, 5min) at 4°C, washed with 70% EtOH, dried and resuspended in 300µl of prehybridisation solution (5 X SSPE, 50% (v/v) formamide, 5% (w/v) dextran sulphate, 1 X Denhardt's solution (2mg/ml BSA, 2mg/ml Ficoll, 2mg/ml

polyvinylpyrrolidone), 250µg/ml of yeast tRNA, 500µg/ml sheared salmon sperm DNA) and used for in situ hybridisation without further dilution.

To evaluate the labelling reaction, 5µl of 1xSSC (0.15M NaCl, 10mM NaH<sub>2</sub>PO<sub>4</sub>, 1mM EDTA, pH7.4) was added to 1µl of probe and the mixture boiled for 5 min, cooled on ice and centrifuged at room temperature for 5 min at 12,000g. 1-2µl samples of the mixture were spotted onto a small nitrocellulose strip, and baked between two sheets of Whatman 3MM paper at 80°C for 30min. The filter was soaked in 2 X SSPE, washed twice for 5 min in PBT (1 X PBS, 0.2% (w/v) BSA Fraction V, 0.1% (v/v) Triton X-100<sup>®</sup>), and blocked for a further 30 min in PBT. The filter was incubated for 30-60 min with the anti-digoxigenin antibody conjugated with alkaline phosphatase (Boehringer Mannheim) diluted at 1:2000 in PBT for 30-60 min, before being washed 4 times for 15 min in PBT with a final wash of 15 min in 100mM NaCl, 50mM MgCl<sub>2</sub>, 100mM Tris pH9.5, 0.1% (v/v) Tween<sup>®</sup> 20. The filter was developed with NTB and BCIP X-phosphate as described by the manufacturers (Boehringer Mannheim).

## Sequencing

Sanger dideoxy sequencing (Sanger *et al.*, 1977) was carried out on denatured double stranded plasmid DNA, using <sup>35</sup>S labelled dATP, a Sequenase kit (version 2.0, United States Biochemical) and the manufacturers recommended protocol without modification. Sequencing reactions were separated by electrophoresis as described elsewhere. Sequences were used to search the NCBI database using the BlastN and BlastX algorithms.

## **2.4.5 Size Fractionation of nucleic acids**

### **Agarose Gel Electrophoresis**

DNA was size fractionated on agarose gels of varying concentration (0.8-1.5% (w/v)) depending on the size of fragments to be resolved. Gels were made and run in 1 X TBE (90mM Tris-borate, pH 8.3, 2mM EDTA) with an applied voltage in the range 2-10 V/cm. Fragment sizes were estimated by comparison to commercial DNA markers (123bp or 1kb ladder, BRL). Gels were stained in EtBR solution (0.5µg/ml), and washed in distilled H<sub>2</sub>O prior to visualisation of DNA fragments by induced UV fluorescence (254nm or 302nm). A photographic record of gels was made on 545- or 667- land film using a polaroid camera fitted with a Kodak Wratten filter No. 23A.

### **Denaturing agarose Gel electrophoresis of RNA**

RNA was size fractionated on 1% (w/v) agarose formaldehyde gels (Sambrook *et al.*, 1989) made in 1 X MOPS (40mM MOPS, pH7.0, 10mM sodium acetate, 1mM EDTA, 2.2M formaldehyde), and run in the same buffer with constant circulation. RNA samples, in a volume of 5µl, were denatured by the addition of 10µl formamide, 2µl 5 X MOPS, 3.5µl formaldehyde, 1µl of EtBr (1µg/ml), and heating to 65°C for 15min., prior to the addition of 2µl loading buffer (30% (w/v) Ficoll 400, 1mM EDTA, 0.25% (w/v) bromophenol blue, 0.25%(w/v) xylene cyanol) and loading onto the gel. Gels were photographed as described above.

### **Denaturing gel electrophoresis of sequencing reaction products**

Products of DNA sequencing reactions (as described elsewhere) were resolved on denaturing polyacrylamide gels (6% (w/v) acrylamide (acrylamide: N,N'-bisacrylamide, 18:1), 6M urea, in 1 X TBE). To 150ml of the acrylamide/urea solution

was added 20 $\mu$ l of TEMED (N,N,N',N'-tetramethylethylenediamine) and 1ml of freshly prepared 10% (w/v) ammonium persulphate to initiate polymerisation, which was allowed to proceed overnight. Gels were pre-run at 2500V for 1hr. before use. Prior to loading on the gel samples were denatured at 72°C for 2 min. in 1 X loading buffer (as supplied in Sequenase Version 2.0 kit). Gels were run for 2-5hr at 2500V and 50°C, then dried onto Whatman 3MM paper at 80°C under vacuum. Autoradiography was carried out without intensifying screens at room temperature.

### **Visualisation and Photography of gels**

DNA and RNA were visualised by UV induced fluorescence on either a short wave (254nm) or long wave (302nm) transilluminator after staining agarose gels with EtBr or adding EtBr to the sample (0.5 $\mu$ g/ml). Gels were photographed using a Polaroid camera loaded with 545- or 667-land film and fitted with a Kodak Wratten filter no. 23A.

### **2.4.6 Blotting and nucleic acid hybridisation**

#### **Southern blotting and DNA/DNA hybridisation**

DNA within agarose gels was transferred to nylon membranes (Hybond-N) by capillary action using essentially the method of Sambrook *et al.*, (1989), and fixed using UV treatment. Filters were pre-hybridised in hybridisation solution (5 X SSPE, 10 X Denhardt's solution, 1% SDS, 0.005% (w/v) sodium pyrophosphate, and 100 $\mu$ g/ml of denatured sonicated salmon sperm DNA) for 1hr at 65°C prior to addition of the radioactive probe. Hybridisation was continued for a minimum of 16hr before filters were washed in 2 X SSPE, 0.1% SDS for 15min., 1 X SSPE, 0.1% SDS for 15 min., and a final wash in 0.1 X SSPE, 0.1% SDS for 15min. (all washes at 65°C). Filters were dried, covered in Saran Wrap<sup>TM</sup> and autoradiographed.

## **Northern blotting and DNA/RNA hybridisations**

Agarose formaldehyde gels containing RNA were transferred to reinforced nitrocellulose (Hybond C<sup>+</sup>), by capillary action (as recommended by Amersham). RNA was fixed to the membrane by baking for 2hr at 80°C. DNA/RNA hybridisations were carried out at 42°C in 50% (v/v) formamide, 5 X SSPE, 2 X Denhardt's solution and 0.1% SDS. Filters were pre-hybridised at 42°C for at 3hr before addition of the denatured radioactive probe ( $10^5$ - $10^6$  cpm/ml of hybridisation solution) and hybridised for a minimum of 16hr. After hybridisation, the filters were washed in 1 X SSPE, 0.1% SDS at room temperature for 20min and then 3 times in 0.2 X SSPE, 0.1% SDS at 65°C for 20min. Prior to autoradiography, filters were blotted dry and covered in Saran Wrap

### **2.4.7 Autoradiography**

Autoradiography of probed filters was carried out at -70°C, where necessary with intensifying screens and exposed to Fuji NIF RX X-ray film, for as long as required. Films were developed using a Kodak X-Omat film processor.

**Chapter 3**  
**Reverse Northern Analysis of Rescued-Plasmids**

### 3.1 Introduction

As discussed in the introduction twenty four fly lines (a gift from J.D. Armstrong and M. Yang), that have P[GAL4] enhancer-trap expression predominantly in the mushroom bodies of the *Drosophila* brain, were selected for further analysis. Initially, it was decided to determine which of the fly lines had genes immediately adjacent to the P[GAL4] insertion site, by subjecting flanking genomic DNA to reverse Northern analysis. The structure of the P[GAL4] element (the plasmid form of the P[GAL4] element, known as pGawB (Brand and Perrimon, 1993), is shown in figure 3.1) allows the acquisition of genomic DNA adjacent to the P[GAL4] insertion site, by plasmid-rescue. Restriction sites within the P[GAL4] element permit the generation of genomic DNA fragments containing the *amp<sup>R</sup>* gene and origin of replication. Upon circularisation and transformation into *E. coli*, the pertinent genomic fragments can be isolated using ampicillin selection. The method of plasmid-rescue is summarised in figure 3.2. Both genomic DNA from the 5' flanking region and the 3' flanking region can be "rescued" in this manner. For the isolation of 3' flanking DNA the following restriction enzymes can be used: *PstI*, *Sall*, *XhoI*, and *SstI*. However, only a single restriction enzyme, *KpnI*, can be used to isolate 5' flanking DNA. (Note that *BstXI* and *SfiI* are not suitable for plasmid-rescue as their recognition sequences are CCANNNNNTGG and GGCCNNNNNGGCC respectively). All the above mentioned enzymes have a six base pair recognition site and would therefore cut random sequence DNA every 4096bp on average. Hence, it should be possible to isolate up to several kb of flanking genomic DNA using this approach.

### 3.2 Isolation of rescued-plasmids

Plasmids were generated containing 3' flanking genomic DNA for all 24 fly lines, however, only 14 of the fly lines yielded plasmids containing 5' flanking DNA. Presumably the fly lines that did not generate 5' plasmids do not have a *KpnI* restriction

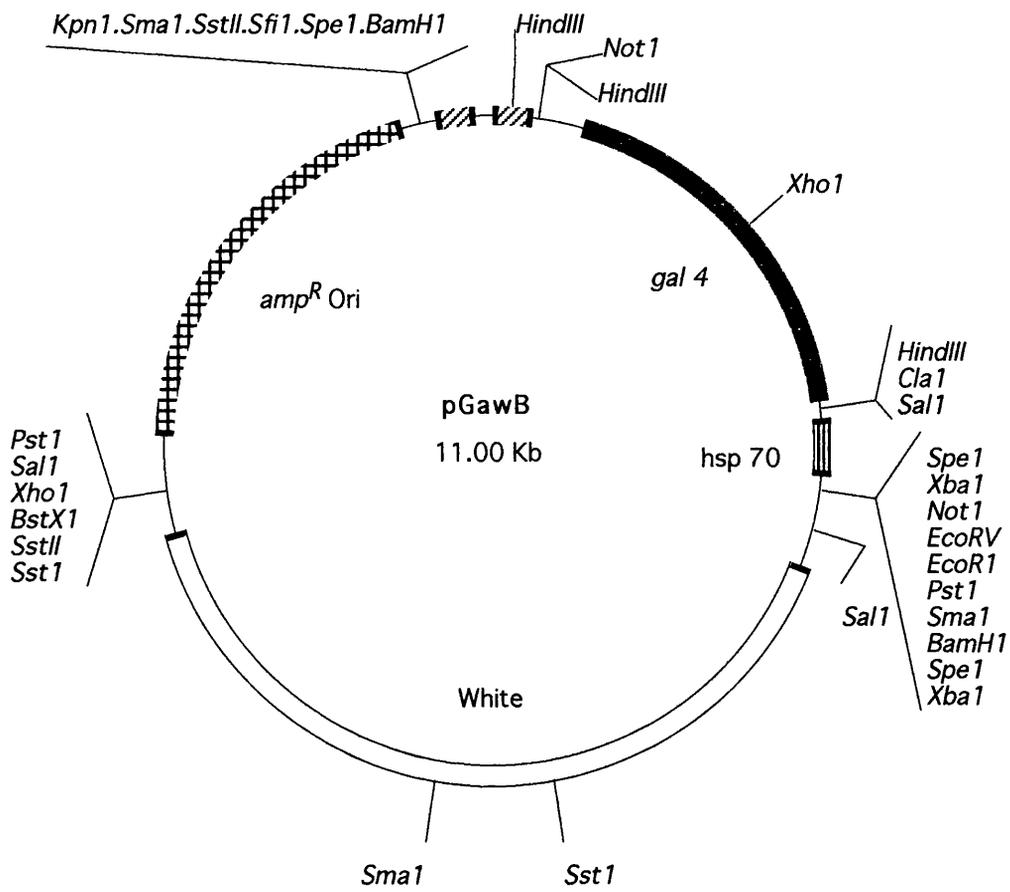
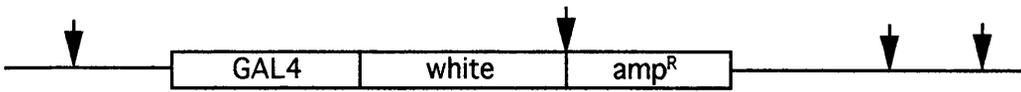
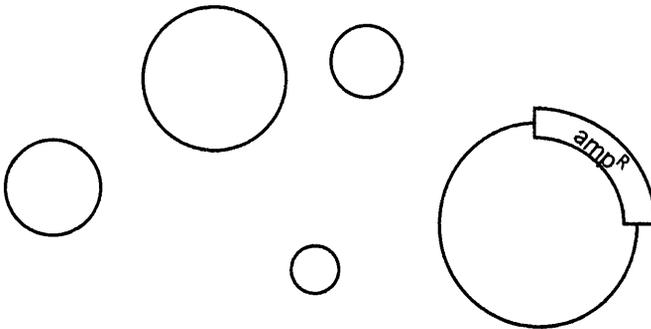


Figure 3.1 Restriction map of the pGawB plasmid, indicating major components and restriction sites of the element.

1. Digest genomic DNA with appropriate enzyme.



2. Dilute Ligation.



3. Electrotransform into *E. coli*, plate with ampicillin selection.

4. Pick colonies, grow overnight culture, prepare plasmid.

**Figure 3.2.** Outline of plasmid-rescue procedure, indicating key steps involved.

site in a suitable location relative to the P[GAL4] insertion site. Figure 3.3 lists the plasmids generated and the enzymes used to generate them. These plasmids contain a total of 167.0kb of genomic DNA, with the 3' plasmids containing on average 3.7kb, and the 5' plasmids 5.6kb, of genomic DNA. To confirm that the rescued-plasmids generated are from the actual P[GAL4] loci, a number of Southern blots were performed. Genomic DNA from the P[GAL4] containing fly lines was digested with the same restriction enzyme used to generate the rescued plasmid from that particular line, and 5µg of the digested DNA loaded onto gels. In lanes adjacent to the digested genomic DNA, 0.1ng of the corresponding linearised rescued-plasmid was also loaded. After electrophoresis gels were blotted onto nylon filters and probed with pBluescript labelled with <sup>32</sup>P. If the rescued-plasmids are from the corresponding P[GAL4] loci then the pBluescript probe should hybridise to a genomic DNA band of the same size as the linearised rescued-plasmid. Results of the Southern blots are shown in figure 3.4. For all fly lines (except 181Y) a genomic DNA band of the same size as the linearised plasmid hybridised to the probe, indicating that generally the rescued-plasmids were generated from the relevant P[GAL4] loci. For fly line 181Y further work generated a new rescued-plasmid of 3.7kb corresponding to the size of genomic DNA band identified.

### 3.3 Reverse Northern analysis

To determine if the rescued-plasmids contained genes or transcriptional units they were subjected to reverse Northern analysis. Furthermore, to determine if any of these plasmids contain transcriptional units that have head elevated or specific expression, the plasmids were probed separately with <sup>32</sup>P labelled cDNA derived from either head polyA<sup>+</sup> RNA or body polyA<sup>+</sup> RNA. Head elevated expression may indicate that the gene is highly expressed in neuronal, and possibly mushroom body, tissue. The 3' plasmids were digested with *SstI* and *BamHI*, and the 5' plasmids with *HindIII*, *SstII*, and *KpnI*. Digested plasmids were loaded onto duplicate gels and restriction fragments separated by electrophoresis. Control plasmids containing the opsin cDNA and an

Fly Line	5' plasmid	3' plasmid
11Y	-	5.0 ( <i>PstI</i> )
17Y	1.8	3.4 ( <i>XhoI</i> )
30Y	4.9	1.5 ( <i>PstI</i> )
43Y	12.5	1.8 ( <i>SstI</i> )
72Y	-	7.5 ( <i>SstI</i> )
103Y	12.0	10.0 ( <i>PstI</i> )
117Y	-	0.7 ( <i>PstI</i> )
119Y	-	3.2 ( <i>PstI</i> )
154Y	-	1.0 ( <i>PstI</i> )
181Y	-	0.7 ( <i>XhoI</i> )
201Y	-	6.0 ( <i>PstI</i> )
238Y	7.9	3.5 ( <i>PstI</i> )
277Y	-	10.0 ( <i>PstI</i> )
C35	8.0	3.5 ( <i>PstI</i> )
C97	5.0	4.0 ( <i>PstI</i> )
C115	-	1.0 ( <i>XhoI</i> )
C245	1.2	3.1 ( <i>PstI</i> )
C302	7.0	2.5 ( <i>PstI</i> )
C532	-	3.2 ( <i>SstI</i> )
C739	2.8	0.7 ( <i>SstI</i> )
C747	2.0	1.2 ( <i>PstI</i> )
C758	2.0	2.3 ( <i>SstI</i> )
C831	5.1	3.5 ( <i>SstI</i> )
75	5.8	9.7 ( <i>SstI</i> )

Figure 3.3. List of rescued-plasmids. Extent of rescued genomic DNA is given in kb. Enzymes used to generate 3' plasmids are given for each fly line. Note that all 5' plasmids were generated using *KpnI*.

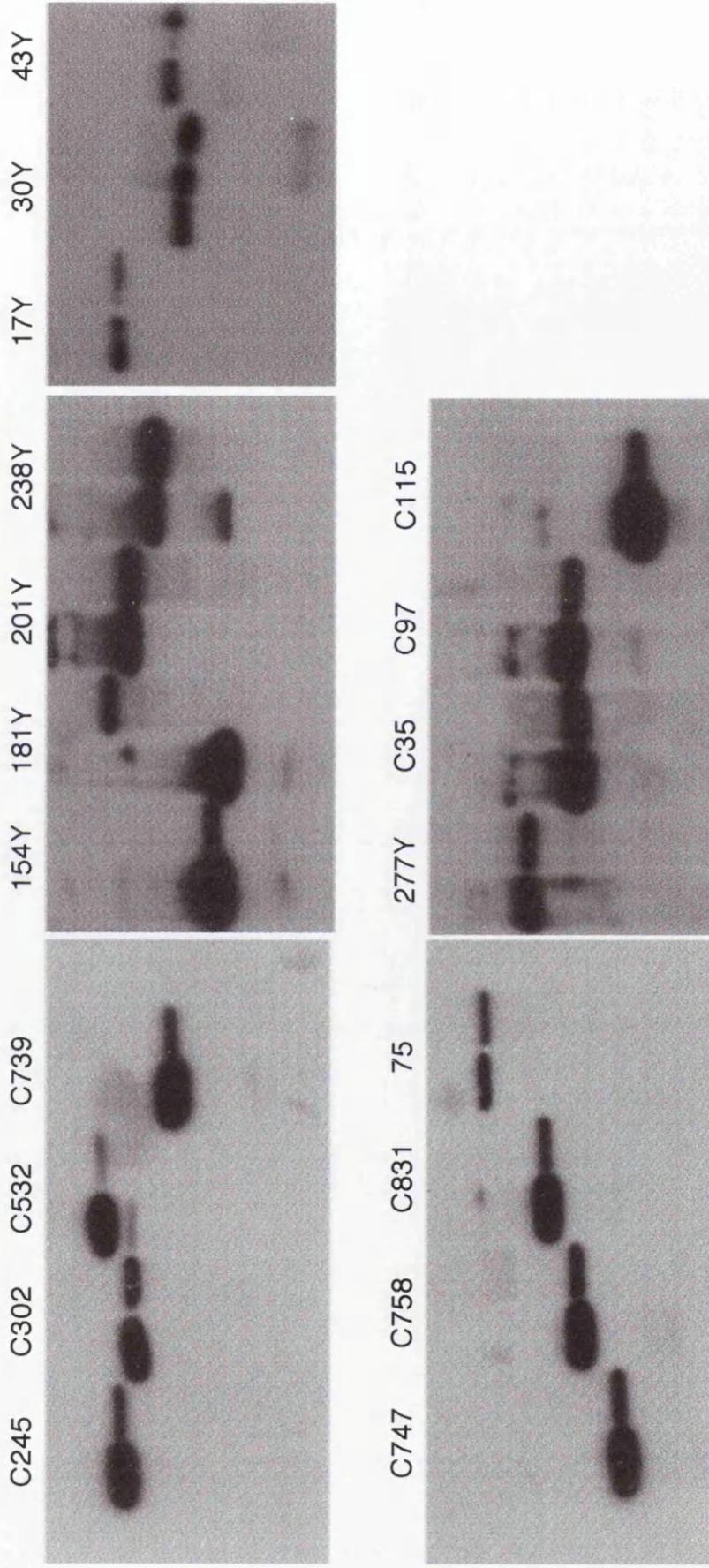


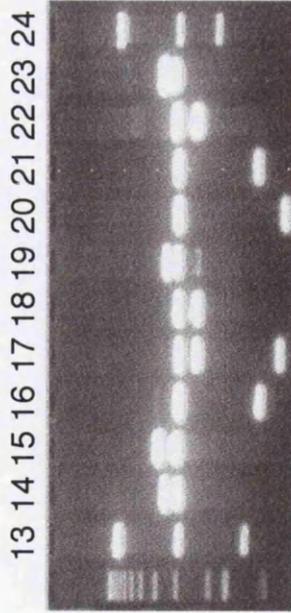
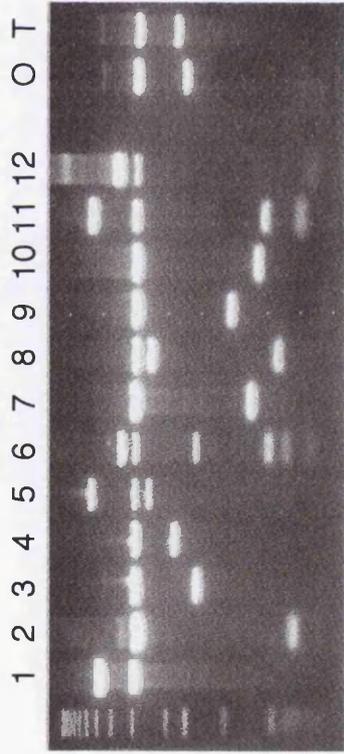
Figure 3.4. Southern blots comparing rescued plasmids to genomic DNA from indicated fly lines, probed with pBluescript. For each pair, genomic DNA is on the right, linearised plasmid on the left. Genomic DNA and plasmids were digested with the restriction enzyme used to generate the particular rescued plasmid.

$\alpha$ -tubulin cDNA were also included on the gels (figure 3.5). The 3' plasmids all have a common 3kb band, derived from the P[GAL4] element, that contains the *amp<sup>R</sup>* gene and the origin of replication. The 5' plasmids all have three P[GAL4] derived bands in common; a 5kb band containing the *white* gene, and two 3kb bands, one containing the *Gal4* gene, and the other the *amp<sup>R</sup>* gene and ori. Gels containing digested plasmids were blotted onto nylon filters, probed as described above, and exposed to X-ray film for 93hrs at -70°C. The results are shown in figures 3.6 to 3.9.

First it must be noted that the specificity of the probes is good, as the head probe hybridises very strongly to the *opsin* cDNA, whilst the body probe only hybridises very weakly to the same cDNA. Both probes hybridise with approximately equal intensity to the ubiquitously expressed  $\alpha$ -tubulin cDNA. The 3' and 5' plasmids all display a 3kb band that appears to hybridise specifically to the head probe. This band contains the *amp<sup>R</sup>* gene of the P[GAL4] element. For reasons that are not entirely clear sequences in the head probe appear to hybridise specifically to the *amp<sup>R</sup>* gene. This "spurious" hybridisation has been observed by others (Milligan, 1995), and is ignored in the analysis presented here. Furthermore, all the 5' plasmids also contain a 5kb band that displays weak specific hybridisation to the head probe. This band contains the *white* gene of the P[GAL4] element and therefore would be expected to hybridise to the head probe. This hybridisation is also ignored in the analysis presented here.

Many of the rescued plasmids appear to contain transcriptional units. A subjective assessment of the levels of expression for these transcriptional units is contained in figure 3.10. An initial examination of the data summarised in figure 3.10 reveals a discrepancy, i.e. nineteen of the twenty four 3' plasmids contain transcriptional units, whilst only four of the fourteen 5' plasmids contain detectable transcriptional units. This discrepancy is even greater when the amount of genomic DNA contained within the two classes of plasmid is considered, i.e. there are nineteen transcriptional units in 89.0kb of 3' flanking DNA, but only four transcriptional units in 78.0kb of 5' flanking DNA. There is, however, a possible simple explanation. The

### 3' PLASMIDS



### 5' PLASMIDS

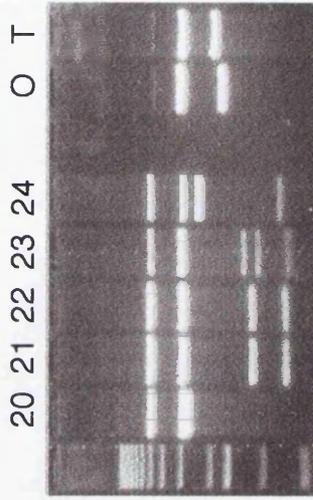
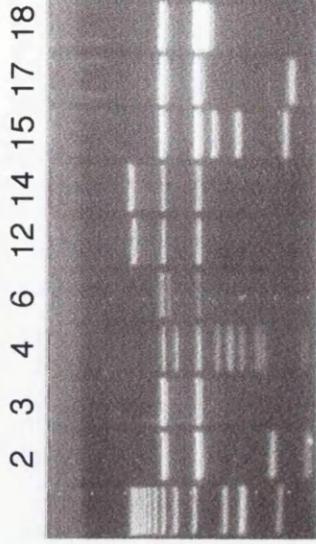
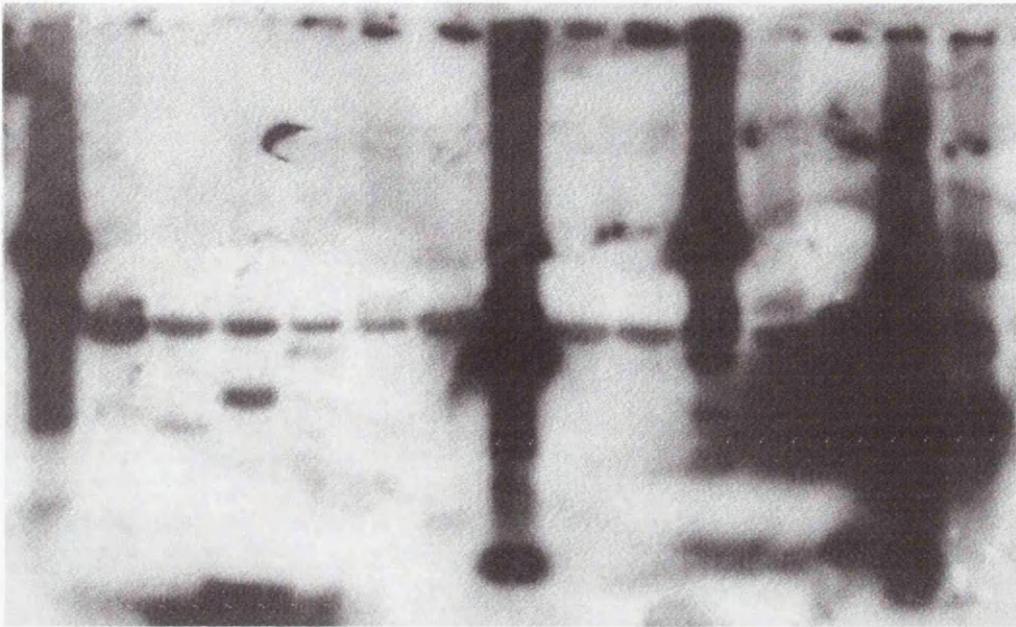


Figure 3.5. Rescued plasmids used for reverse Northern analysis. 3' plasmids were digested with *SstI* + *BamHI*, 5' plasmids were digested with *HindIII* + *SstII* + *KpnI*. Lane numbers correspond to fly lines as follows: 1 = 11Y, 2 = 17Y, 3 = 30Y, 4 = 43Y, 5 = 72Y, 6 = 103Y, 7 = 117Y, 8 = 119Y, 9 = 154Y, 10 = 181Y, 11 = 201Y, 12 = 238Y, 13 = 277Y, 14 = C35, 15 = C97, 16 = C115, 17 = C245, 18 = C302, 19 = C532, 20 = C739, 21 = C747, 22 = C758, 23 = C831, 24 = 75. Opsin (O) and  $\alpha$ -Tubulin (T) controls are also included.

HEAD PROBE

1 2 3 4 5 6 7 8 9 10 11 12 O T



BODY PROBE

1 2 3 4 5 6 7 8 9 10 11 12 O T

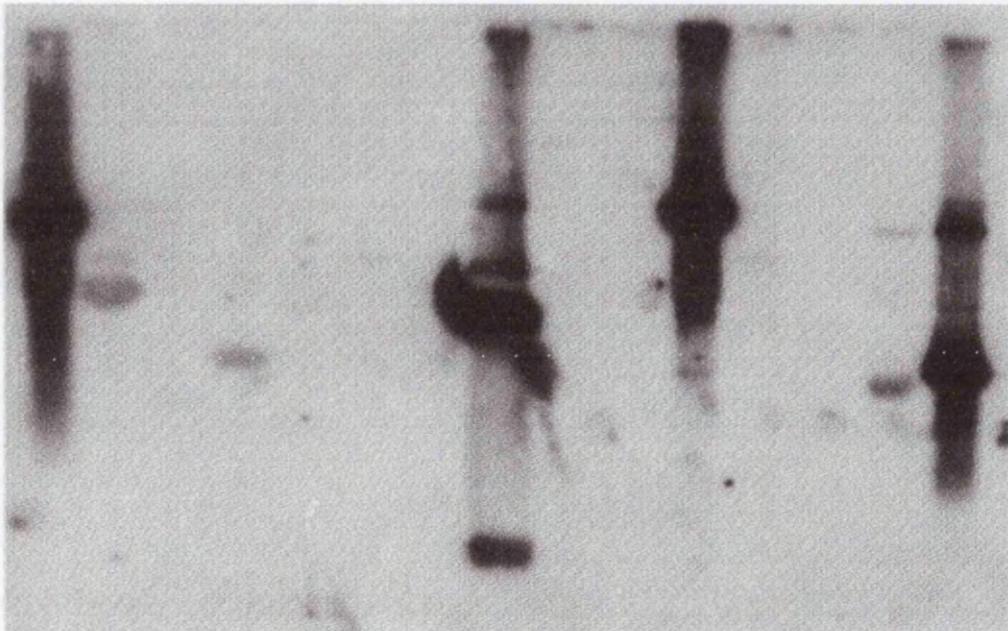


Figure 3.6. Reverse Northern analysis of 3' rescued plasmids. Lane numbers correspond to those used in figure 3.5.

HEAD PROBE

13 14 15 16 17 18 19 20 21 22 23 24



BODY PROBE

13 14 15 16 17 18 19 20 21 22 23 24

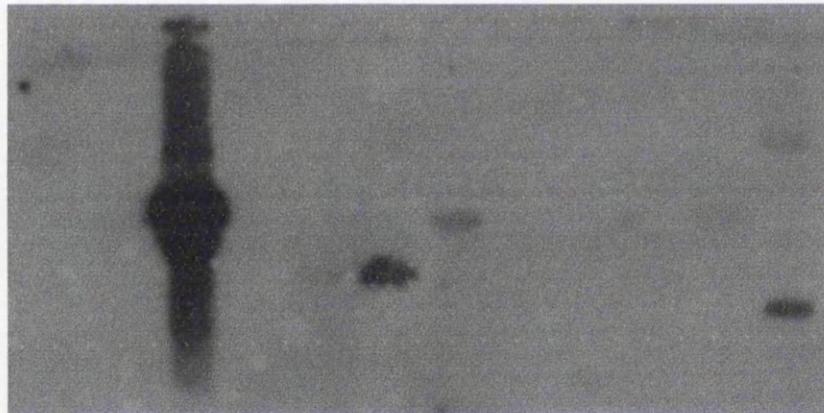
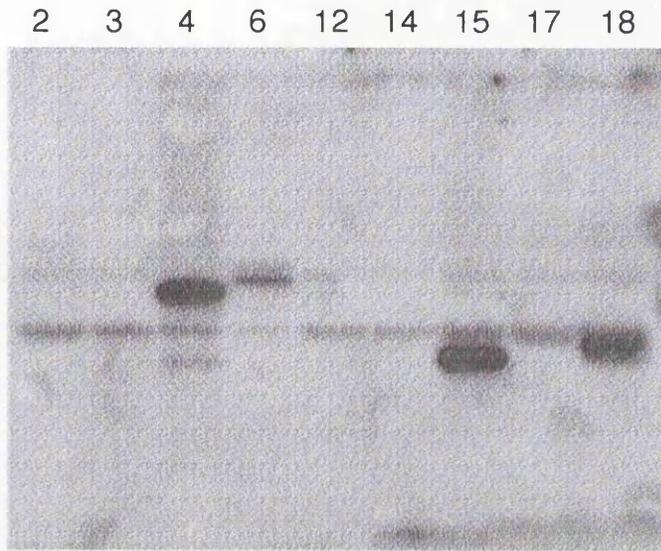


Figure 3.7. Reverse Northern analysis of further 3' rescued plasmids. Lane numbers correspond to those used in figure 3.5.

HEAD PROBE



BODY PROBE

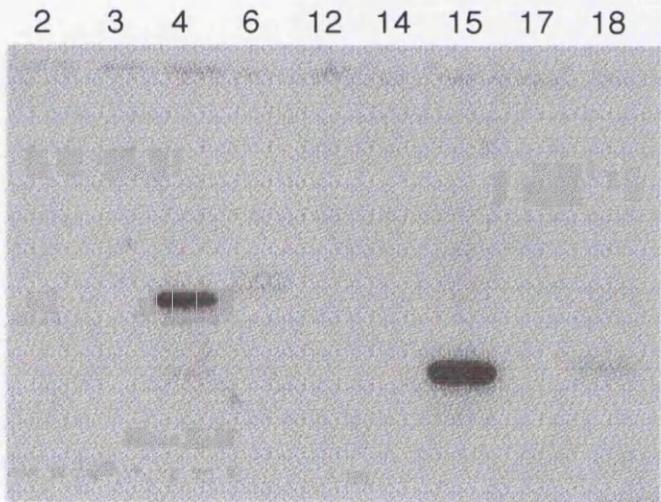


Figure 3.8. Reverse Northern analysis of 5' plasmids. Lane numbers correspond to those used in figure 3.5

HEAD PROBE

20 21 22 23 24 O T



BODY PROBE

20 21 22 23 24 O T



Figure 3.9. Reverse Northern analysis of further 5' plasmids. Lane numbers correspond to those used in figure 3.5.

Fly Lines	3' Plasmid		5' Plasmid	
	Head Probe	Body Probe	Head Probe	Body Probe
11Y	+++	+++	?	?
17Y	++	++	-	-
30Y	+	+	-	-
43Y	++	+	++	++
72Y	+	-	?	?
103Y	-	-	+	+
117Y	+	+	?	?
119Y	+++	+++	?	?
154Y	-	-	?	?
181Y	-	-	?	?
201Y	+++	+++	?	?
238Y	+	-	-	-
277Y	+	-	?	?
C35	-	-	-	-
C97	+++	+++	-	-
C115	+	-	?	?
C245	+	+	-	-
C302	++	++	++	+
C532	+	+	?	?
C739	-	-	-	-
C747	+	+	-	-
C758	+	+	-	-
C831	+	+	-	-
75	+	+	-	-

Figure 3.10. Relative expression of genes contained in rescued-plasmids, based on results of reverse Northern shown in figures 3.6. to 3.9. +++ = strong expression, ++ = medium expression, + = weak expression, - = no visible expression, ? = plasmid not available.

rescued-plasmids containing 5' flanking DNA carry 11 kb of the P[GAL4] element, compared to only 3 kb of P[GAL4] element contained by the 3' rescued-plasmids. Therefore, each microgram of 5' plasmid will, on average, contain less genomic DNA than a microgram of 3' plasmid. For all of the rescued plasmids approximately 1 $\mu$ g was loaded onto the gels. Therefore, the reverse Northern analysis of the 5' plasmids will have been less sensitive than the analysis of the 3' plasmids, and weakly expressed transcriptional units, within the 5' plasmids, may not have been detected.

### 3.4 Discussion

The reverse Northern analysis, outlined in the results section, identified many of the rescued-plasmids as containing transcriptional units. It was clearly not feasible to perform the detailed molecular studies necessary to identify the transcriptional units on all the rescued-plasmids. Therefore, it was decided to select two fly lines (and their associated rescued-plasmids) for further molecular work. One fly line chosen for further study was 201Y. A 3' plasmid had been generated for this line containing a transcriptional unit that is strongly expressed at similar levels in heads and bodies. This fly line has been the subject of extensive studies by others, in relation to mushroom body structure, and also as a tool towards understanding fly sexual behaviour (Yang *et al.*, 1995, O'Dell *et al.*, 1995, see Introduction for a fuller description of these studies). It was therefore thought pertinent that these studies be complemented by a molecular analysis of the 201Y P[GAL4] locus.

Generally, the purpose of this study is to clone and identify genes flanking the P[GAL4] insertions, and to determine if any of these genes have expression that reflects the adjacent enhancer-trap expression within the mushroom bodies. Clearly, those fly lines whose rescued-plasmids do not contain transcriptional units, as determined by reverse Northern analysis, require further work (e.g. the isolation of genomic DNA clones) before genes close to the P[GAL4] element can be identified. More rapid progress can be made with fly lines whose rescued-plasmids contain

transcriptional units. This immediately ruled out fly lines 154Y, 181Y, C35, and C739, as possible candidates for further study, as none of their associated rescued-plasmids have detectable transcriptional units.

The reverse Northern analysis, discussed earlier, was performed with cDNA probes copied from fly head or body mRNA. This was done explicitly to determine if any of the rescued-plasmids contain transcriptional units with head elevated or specific expression. Although fly bodies contain neuronal tissue (particularly within the thoracic ganglion), fly heads are, by comparison, relatively enriched for neuronal tissue. Therefore, it is not an unreasonable assumption that head elevated expression may well indicate neuronal expression. Of course, this does not necessarily imply specific or elevated expression within the mushroom bodies. Nevertheless, head elevated expression is still a useful criterion for selecting fly lines for further studies. Six fly lines, 43Y, 72Y, 238Y, 277Y, C115, and C302, have rescued-plasmids that contain transcriptional units with head elevated expression. Of these only two, 43Y and C302, have identified transcriptional units within both 3' and 5' rescued plasmids. The transcriptional unit within the 43Y 3' plasmid is clearly head elevated, as is the transcriptional unit within the C302 5' plasmid. These fly lines are both good candidates for further study, and it was decided that 43Y would be the other fly line, together with 201Y, chosen for further study. Subsequent chapters outline the results of these studies.

**Chapter 4**  
**The 201Y 3' Gene**

## 4.1 Introduction

As discussed in Chapter 3 fly line 201Y was selected for further study partly because it was already the subject of study by other members of the group, and a molecular analysis of the 201Y P[GAL4] genomic locus would complement their work. The 201Y P[GAL4] enhancer-trap is inserted at cytological location 56D (ZonSheng Wang, pers.comm.), and of all the 24 mushroom body expressing enhancer-trap lines discussed in the Introduction and Chapter 3, 201Y has the most mushroom body restricted expression pattern within the CNS (Armstrong, 1995). Figures 4.1 and 4.2 show the GAL4 mediated *lacZ* expression pattern in adult female flies of the 201Y enhancer-trap. As can be seen, head expression is confined largely to the mushroom bodies of the CNS, while expression in the body of the fly appears to be absent from the thorax, but is present in a few abdominal structures. Plasmid-rescue had generated a 3' plasmid for fly line 201Y. Reverse Northern analysis revealed that this plasmid contained a transcriptional unit that was strongly expressed in both fly heads and bodies. This chapter describes the identification and analysis of this transcriptional unit.

## 4.2 Analysis of the 201Y 3' gene

### 4.2.1 Reverse Northern analysis of pB201Y

Plasmid rescue of genomic DNA flanking the 3' region of the P[GAL4] insertion point, digested with restriction enzyme *PstI*, generated a 9.1kb plasmid known as pB201Y. To facilitate manipulation of this plasmid and further work with it, a restriction map was constructed. The following single and double restriction digests were performed on pB201Y: *PstI*, *BamHI*, *Sall*, *SstI*, *XhoI*, *PstI* and *BamHI*, *PstI* and *EcoRI*, *BamHI* and *EcoRI*, *BamHI* and *Sall*, *BamHI* and *XhoI*, and also the multiple digest *BamHI*, *PstI* and *EcoRI*. Fragments were separated by gel electrophoresis and results are



Figure 4.1. 201Y enhancer-trap expression. Frontal section of a fly head in the plane of the mushroom body lobes. GAL4 mediated *lacZ* expression can be seen in the  $\alpha$ -lobe and  $\beta/\gamma$  lobes.

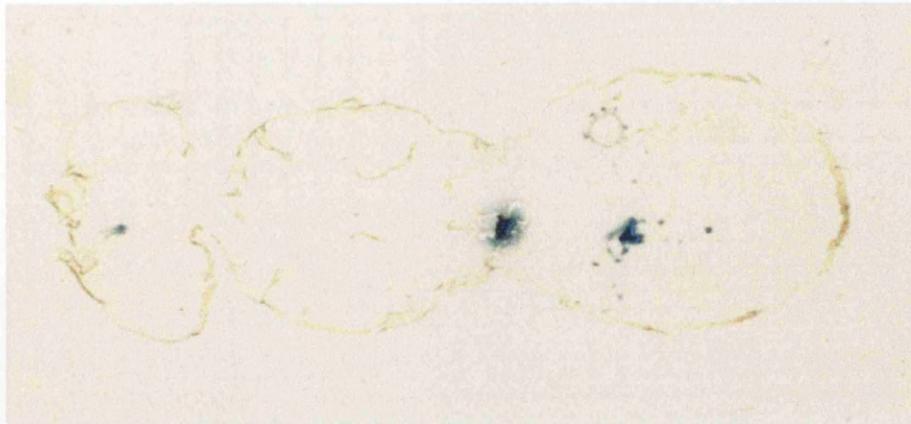


Figure 4.2 Horizontal and vertical sections through a whole 201Y adult female fly, stained for  $\beta$ -galactosidase activity. Staining can be seen in the head, in cells surrounding the gut lumen, in the egg laying apparatus, and in an unidentified structure at the junction between the thorax and abdomen.

shown in figure 4.3. The information contained there, together with the known restriction sites flanking the ampicillin resistance gene of the P[GAL4] element, allows the restriction map depicted in figure 4.4 to be constructed.

As described in Chapter 3, reverse Northern analysis had identified pB201Y as containing a transcriptional unit. Following the construction of a restriction map for pB201Y, it was decided to refine the reverse Northern analysis by repeating it on particular restriction digests of pB201Y. Therefore the following double digests were performed on pB201Y: *SstI* and *EcoRI*, *SstI* and *XhoI*, *SstI* and *BamHI*. Digested DNA was loaded onto duplicate gels, together with the control plasmids containing the cDNAs  $\alpha$ -tubulin and opsin, and restriction fragments separated by electrophoresis. Gels were blotted onto nylon filters, and again probed separately with <sup>32</sup>P labelled single-stranded cDNA, copied from head or body mRNA, to identify any possible transcriptional units. Use of separate head and body cDNA probes would determine whether any identified transcriptional units had head elevated or specific expression. Filters were exposed to X-ray film for 6hrs at -70°C and the results are shown in figure 4.5. The first thing to note is that again the specificity of the cDNA probes was good, with the head probe hybridising very strongly to the opsin cDNA, whilst hybridisation of the body probe to opsin was weak. Hybridisation of both probes to the  $\alpha$ -tubulin cDNA is similar, indicating that they were of a similar specific activity.

Initial comparison of the two reverse Northern, indicates the existence of restriction fragments that hybridise to the head cDNA probe but do not hybridise to the body cDNA probe. However, these particular fragments, i.e. the 3.3kb *SstI/EcoRI* (note that this restriction digest yields two different 3.3kb fragments), the 3.4kb *SstI/XhoI*, and the 2.9kb *SstI/BamHI* restriction fragments, that specifically hybridise to the head cDNA probe, contain the ampicillin resistance gene from the P[GAL4] element. As discussed in Chapter 3, for reasons that are not entirely clear, the ampicillin resistance gene appears to hybridise specifically to the head cDNA probe. In this case the hybridisation to the head cDNA probe is particularly clear and distinct. Nevertheless, it was assumed that the head specific hybridisation seen here is due to

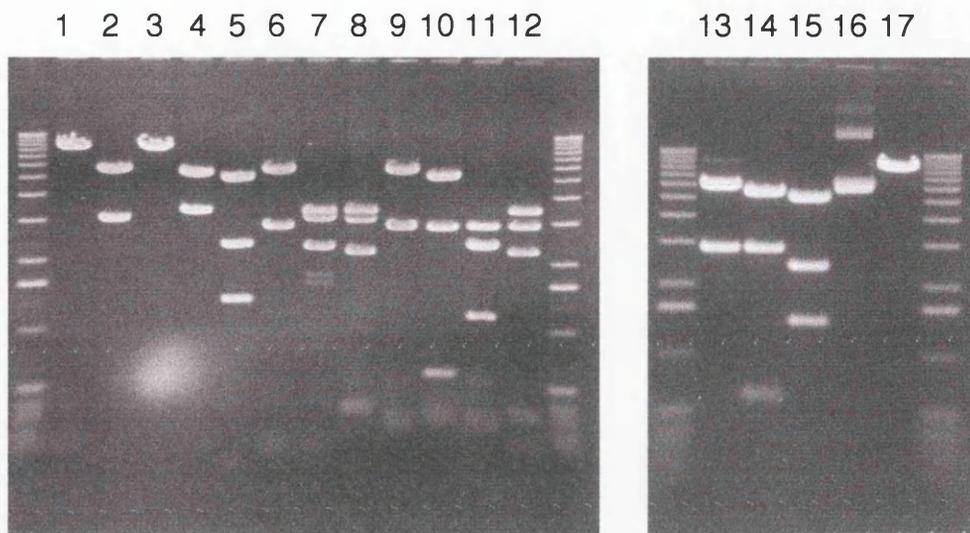


Figure 4.3. Restriction digests of plasmid pB201Y. Fragments sized using BRL 1kb ladder. Lane numbers correspond to the following digests:

1=*PstI*, 2=*BamHI*, 3=*Sall*, 4=*SstI*, 5=*XhoI*, 6=*PstI* + *BamHI*, 7=*PstI* + *EcoRI*, 8=*BamHI* + *EcoRI*, 9=*BamHI* + *Sall*, 10=*BamHI* + *SstI*, 11=*BamHI* + *XhoI*, 12=*BamHI* + *PstI* + *EcoRI*, 13=*SstII*, 14=*SstII* + *SstI*, 15=*HindIII* + *XhoI*, 16=*HindIII*, 17=*HindIII* + *PstI*.

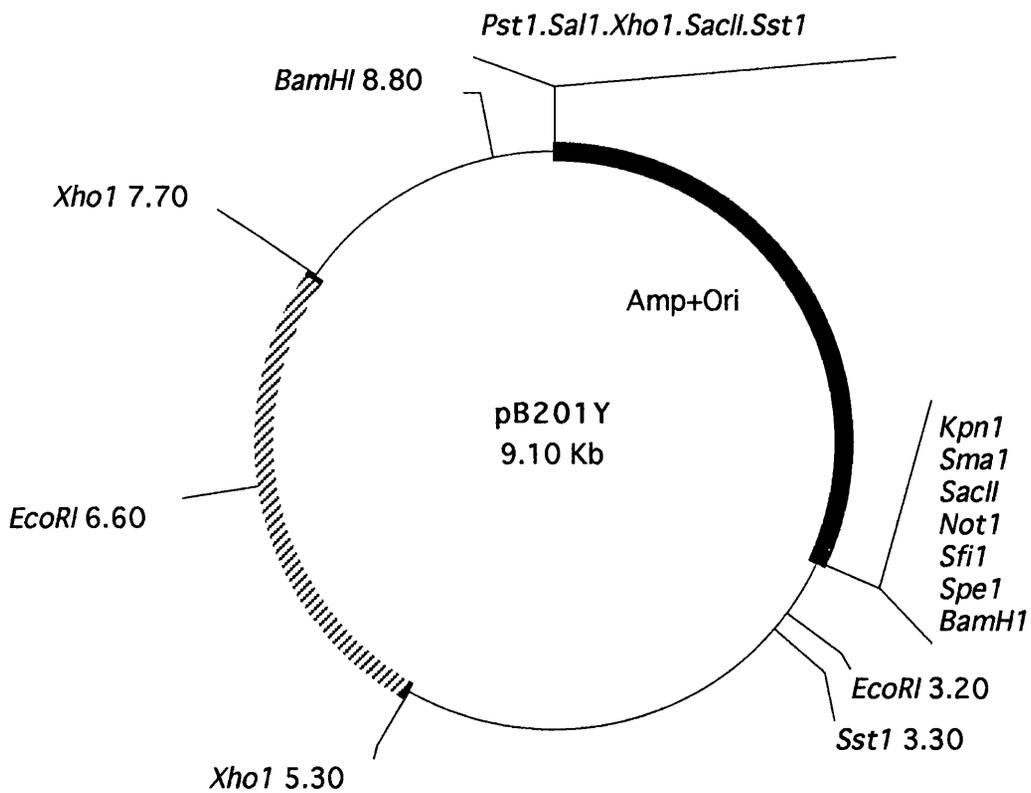


Figure 4.4. Restriction map of plasmid pB201Y. Numbers indicate distance in kb from "twelve o'clock" position. Thick black line is ampicillin resistance gene from P[GAL4] element. Hatched segment indicates region identified by reverse Northern as containing a transcriptional unit.

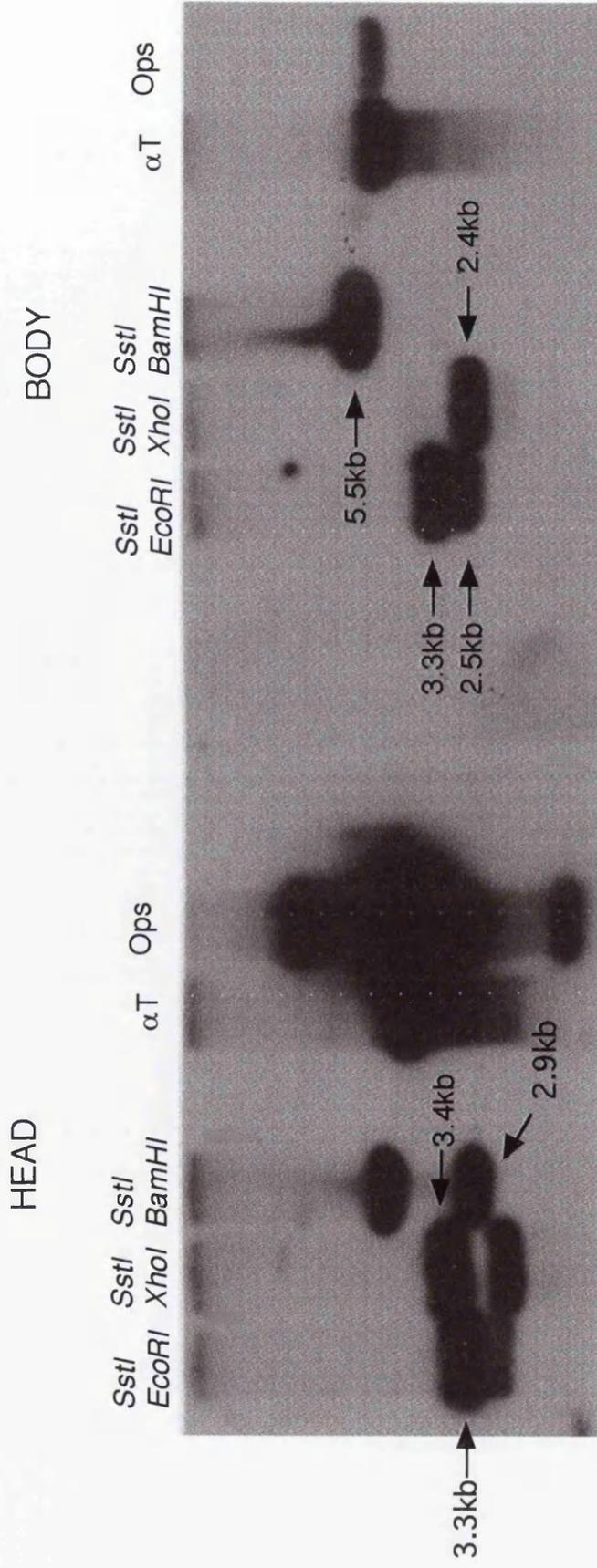


Figure 4.5. Reverse Northern analysis of pB201Y. Restriction digests of pB201Y are indicated above the lanes.  $\alpha$ -Tubulin ( $\alpha$ T) and opsin (Ops) controls are indicated. Sizes of bands discussed in the text are indicated.

this "spurious" hybridisation to the ampicillin resistance gene, and does not represent the existence of a head specific transcriptional unit in the plasmid pB201Y. Therefore, ignoring this "spurious" hybridisation, the following restriction fragments appear to hybridise equally to both probes: the 2.5kb and 3.3kb *SstI/EcoRI* fragments, the 2.4kb *SstI/XhoI* fragment, and the 5.5kb *SstI/BamHI* fragment. All fragments hybridise with equal intensity to both probes indicating that the transcriptional unit identified does not have head elevated or specific expression. Analysis of the hybridising restriction fragments limits the transcriptional unit to within the two *XhoI* sites at 5.30 and 7.70 on the pB201Y restriction map, as indicated on figure 4.2. Furthermore, if there is only one transcriptional unit it must straddle the *EcoRI* site at 6.60, as both the 2.4kb and 3.3kb *EcoRI/SstI* fragments (which together contain the two *XhoI* sites at 5.30 and 7.70) hybridise to the cDNA probes.

#### **4.2.2 Isolation and Sequence Analysis of cDNA clones**

To identify the transcriptional unit or units involved it was decided to screen a *Drosophila* head cDNA library constructed in lambda phage NM1149. Initially 50 000 recombinant phage were plated out and duplicate filter lifts probed with the 2.4kb *XhoI* fragment of pB201Y labelled with <sup>32</sup>P. Two positives were identified (known as 216 and 311) and purified down to single plaques. Phage DNA was prepared from both positives and digested with restriction enzymes *EcoRI* and *HindIII* to release the cDNA insert. Restriction fragments were separated by gel electrophoresis, Southern blotted onto nylon filters, and reprobbed, with the 2.4kb *XhoI* fragment of pB201Y labelled with <sup>32</sup>P, to identify the positives as real. Results are shown in figure 4.6. Each cDNA is approximately 300bp in size with 216 being slightly larger than 311. Both cDNAs hybridised to the pB201Y probe confirming them as real, and were subsequently cloned into the vector pBluescript, giving the new plasmids pB216 and pB311. The two subcloned cDNAs were sequenced using primers T3 and T7 which hybridise to sequences either side of the pBluescript multiple cloning site. For each

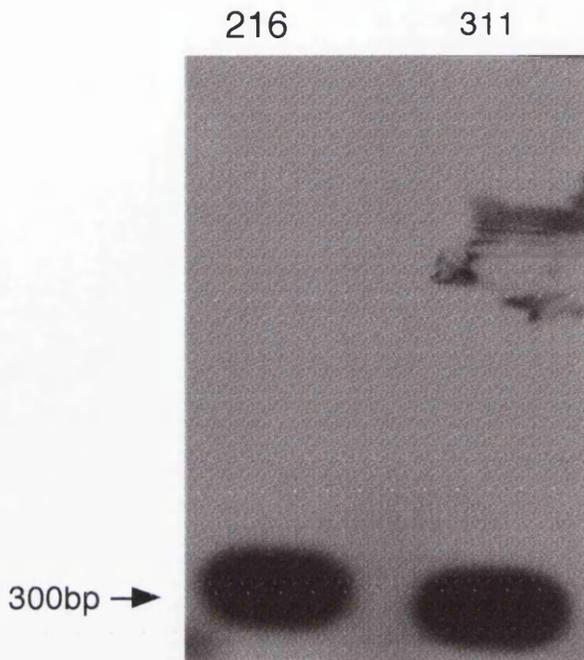


Figure 4.6. cDNAs 216 and 311 digested with *EcoRI* and *HindIII*, probed with the 2.4kb *XhoI* fragment of pB201Y.

plasmid the combined sequences, generated from the T3 and T7 primers, spanned the whole cDNA. Comparison of the sequences obtained indicated a region of overlap between the pB216 and pB311 cDNAs as shown in figure 4.7. The two cDNAs give a single contiguous sequence of approximately 550bp, with the pB311 sequence containing an obvious poly A tail indicating it to be the 3' end of the combined cDNA. Initially there were a few sequence discrepancies within the region of overlap but these were eliminated by sequencing both cDNAs completely on both DNA strands.

The construction of the *Drosophila* head cDNA library, from which the clones 216 and 311 were obtained, involved first-strand cDNA synthesis from poly A<sup>+</sup> mRNA using an oligo dT primer. The question arises as to how the 216 cDNA was generated, as it has no poly A tail nor any obvious site from which oligo dT could misprime? It must represent an artefact or possibly a deleted cDNA. However, as it has a distinct overlap with 311 which has a obvious poly A tail, it was decided to regard the combined sequence as one contiguous cDNA.

A search of the DNA sequence databases was conducted with the combined 216/311 sequence using the BlastN algorithm. Significant homology was found to a class of genes that encode a family of 12kDa proteins known as peptidylprolyl *cis-trans* isomerases and also as FK506 binding proteins (FKBP12), for which no *Drosophila* homologue had yet been identified. An example of the sequence alignments found is given in figure 4.8 for the human FKBP gene sequence and the combined 216/311 DNA sequence. Examination of the 216/311 DNA reveals an open reading frame from base 1 terminating at a stop codon at base 316 This ORF was translated in frame 1, and a comparison between the derived protein sequence and the complete human FKBP12 protein sequence is also given in figure 4.8. The protein sequences are extremely similar, being 75% identical over the 105 aa that are compared.

The 216/311 sequence is apparently incomplete, as the ORF does not contain a translation start codon. However, comparison with the human FKBP12 protein sequence suggests that the 216/311 sequence is only missing three amino acids at the amino terminal, and indeed 12kDa FK506 binding proteins from a number of other

```

1 216 Sequence:      50
CAATAGTTCCAATGGCTCTGGTGATGGCAGCACCTATCCCAAGAATGGCCAAAAGGTACGGTCCACTACA

      100
CCGGCACCCCTGGACGATGGACCAAGTTCGATTCGTGCGCGGACCGCAACAAGCCATTCAAGTTCACCATCG

      150
GCAAGGGGAGGTCAATCCGTGGCTGGGATGAGGGAGTTGCCAGTTGAGCGTCGGCCAGAGCGCCAAGCTGA

      200
      250
TTTGCTCGCCGGACTATGCCCTACGGTAGCCGTGGCCACCCCGGGGTCAATCCGCCCAACTCCACCCTCACCT
|||||
ATGCCCTACGGTAGCCGTGGCCACCCCGGGTCAATCCGCCCAACTCCACCCTCACCT

      300
TCGACGTCGAGCTGCTCAAGGTCGAATAGGC
|||||
TCGACGTCGAGCTGCTCAAGGTCGAATAGGCGCACAGGATGCCCAATGTGTATACCCCAAAACCAATCGCAG

      350
GGGGTTGCGGACCGGCTGCACCGGCTGACGAGCGGAAGAACAATCATATATCGGAACCCAGAAAATCCAGCAGCAT

      400
      450
TTGCCAACCCAAAATAATAATGATAACTTCATACAGCTCTAAACTAGAACTAATTTAACCCGCAAAGGGAACG

      500
CATCTTCTACCAATACAAAATAAACATTTATTCAAGTCAAAAAA

```

Figure 4.7. Alignment between sequences from 216 and 311 cDNAs.

```

216/311: 1 CAAGTAGTCCAATTGCTCCTGGTGATGGCAGCACCTATCCCAAGAATGGCCAAAAGGTC 60
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
FKBP12: 119 CAGGTGGAACCCATCTCCCCAGGAGACGGCGCCACCTTCCCCCAAGCGCGCCAGACCTGC 178

216/311: 61 ACGGTCCACTACACCCGGCACCCCTGGACCATGGCACCACCAAGTTCGATTCGTCGGCCGACCCG 120
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
FKBP12: 179 GTGGTGCACACTACACCCGGGATGCTTGAAGATGGAAGAAATTTGATTCCTCCCGGGACAGA 238

216/311: 121 AACAAAGCCATTCAAGTTCACCATCGCAAGGGCGAGGTCAATCCGTGGCTGGGATGAGGGA 180
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
FKBP12: 239 AACAAAGCCCTTAAAGTTATGCTAGGCAAGCAGGAGGTGATCCGAGGCTGGGAAGAAGGG 298

216/311: 181 GTTGCCCGAGTTGAGCGTCGGCCAGAGCCGCAAGCTGATTTGCTCGCCGGACTATGCCCTAC 240
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
FKBP12: 299 GTTGCCCGCAGATGAGTGTGGGTCAGAGAGCCAAACTGACTATATCTCCAGATTATGCCCTAT 358

216/311: 241 GGTAGCCGTGGCCACCCCGCGTCATTCGGCCCAACTCCACCCTCACCCTTCGACCGTCGAG 300
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
FKBP12: 359 GGTGCCACTGGGCACCCAGGCATCATCCCAACCACATGCCACTCTCGTCTTCGATGTGGAG 418

216/311 301 CTGCTCAAGGTCGAAT 316
          ||| ||| ||| ||| |||
FKBP12: 419 CTTCTAAAACCTGGAAT 434

216/311 translated: QVVP IAPGDGSTYPKNGQKVTVHYTGLDDGTFKFDSSRRDRNPKPFKFTIGKGEVIRGWDEGVAQLSVG
Human FKBP12      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
                   MGVQVETISPGDGRTFPKRGQTCVHYTGMLEDGKFKFDSSRRDRNPKPFKFM LGKQEVIRGWEEGVAQMSVG

216/311 translated: QRAKLICSPDYAYGSRGHPGVIPPNSTLTFDVELLKVE
Human FKBP12:     ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
                   QRAKLTI SPDYAYGATGHPGIIPPHATLVFDVELLKLE

```

Figure 4.8. Alignment between 216/311 sequence and human FKBP12 sequence (Genbank acc. no. M80199). Upper alignment is of DNA sequence, lower alignment of protein sequence.

species, such as rabbit and mouse, are, at 108 aa long, three amino acids longer than the 216/311 protein sequence. This suggested that the simplest way to complete the sequence of this gene was to obtain the short amount of additional 5' sequence required from the genomic DNA contained in pB201Y, rather than screen cDNA libraries for full-length cDNAs. Therefore, based on knowledge of the 216/311 DNA sequence a primer was designed to obtain sequence from pB201Y.

Alignment between the genomic sequence obtained from pB201Y and 216/311 cDNA sequence is shown in figure 4.9. The alignment between the cDNA and genomic sequences is not continuous, suggesting that the genomic sequence contains an intron. To check that the discontinuous alignment of the cDNA and genomic sequences is due to the presence of an intron, the splice junction sequences were compared to the consensus sequences for *Drosophila* splice junctions. *Drosophila* splice junctions have been studied in some detail by Mount *et al.*, 1992. Using a database of 209 *Drosophila* introns they arrived at 5' and 3' splice site consensus sequences for both small introns (80bp or less in length) and large introns (greater than 80 bp in length). Figure 4.10 shows the comparison between the putative identified splice sites and the splice site consensus sequences for large introns. There is good agreement between both the 5' and 3' putative splice sites and the relevant consensus sequences. In particular, bases that are invariant or highly conserved in the consensus sequences are present in the genomic sequences. Some of the differences that do occur, e.g. at positions 2 and 4 of the 5' splice site, and positions -12, -5, and 2 of the 3' splice site, are minor deviations from the consensus, as in all these cases the actual base is the second most frequently found. Only at positions 6 of the 3' splice site, and position -8 of the 5' splice site, are there major deviations from the consensus. In both cases a G is present at these positions in the actual sequence, and this occurs at frequencies of 11% and 8% respectively in the examined introns. Nevertheless, overall there is good agreement between the splice site sequences and the determined consensus sequences, suggesting that a real intron has been identified.

```

1                               50
Genomic sequence: CCGTAGCTGTCAATTCGGCGACGGACGAACTGAAATCAACGTGCTTATTTCCGCTCGGTAATTTGCCAGAA

                               100
216/311 sequence: CAAGTAGTTCCAAATTGCTCCTGGTGATG.....
                   |||||||||||||||||||||||||||
Genomic sequence: AAAAACTAATCGAAATGGGGGFACAAGTAGTCCAAATTGCTCCCTGGTGATGGATCCCAATGATTTGGGTT
                   .....
                   150           200
Genomic sequence: .....
                   AGAATCCGCAAGAAATGCCGTAAACTACCGGCTTCTAACCATTATTAGTGCAATTTCCGCAATGCACAATTCG
                   .....
216/311 sequence: .....GCAGCAC
                   .....
Genomic sequence: GAAATACCTGCCACAATATGGTTCGTCGGCTCACACACATCCATTTTGCAATCCCGCCCGCAGGCAGCAC
                   .....
                   300           350
216/311 sequence: CTATCCCCAAGAAATGGCCAAAAGGTCACGGTCCACTACACCCGGCACCCCTGGACGATGGCACCACCAAGTTC
                   ||||
Genomic sequence: CTAT
                   CTGCTACCGTGGTTGAAG
                   Sequencing Primer

```

Figure 4.9. Comparison of pB201Y genomic DNA sequence and 216/311 cDNA sequence. Identity between 216/311 sequence and genomic sequence indicated by vertical bars. Sequence and position of sequencing primer is indicated. Underlined sequence indicates nucleotides that conform to branchpoint consensus WCTAATY (in which branch formation occurs at the highlighted position).

5' Splice site:

	-3	-2	-1	1	2	3	4	5	6	7	
Genomic sequence	A	T	G		G	T	G	T	G	G	A
Consensus sequence	M	A	G		<u>G</u>	<u>T</u>	R	A	<u>G</u>	T	A

3' Splice site:

	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	
Genomic sequence	T	T	C	C	C	C	G	C	C	C	G	C	A	G		G	C
Consensus sequence	T	T	T	Y	Y	Y	T	Y	<u>Y</u>	T	N	C	<u>A</u>	<u>G</u>		R	T

Figure 4.10. Comparison of FKBP12 intron splice sites with the consensus for *Drosophila* large introns (Mount *et al.*, 1992). Consensus sequences use standard IUPAC codes. Vertical lines indicate splice junctions. Double underlined bases are invariant. Single underlined bases are highly conserved (occurring in >75% of cases).

Most *Drosophila* introns are quite small with more than half being less than 80 bp in length, making, at 163 bp, this identified intron significantly larger than the median. But with approximately 10% of *Drosophila* intron lengths falling in the range 100-200 bp, and the largest being several kilobases, this intron is by no means exceptional.

The intron/exon structure of the human FKBP12 gene has been elucidated (DiLella and Craig, 1991). They found the exons of human FKBP12 to correspond to the protein's known structural features, such as  $\beta$ -sheets and  $\alpha$ -helices. Interestingly, the position of the first human FKBP12 intron occurs at exactly the same corresponding position, i.e. within the 13th codon, as the intron described above for the *Drosophila* FKBP12, suggesting FKBP12 intron/exon boundaries are evolutionarily conserved.

Mount *et al.*, (1992), as well as determining splice site consensus, also looked at intron sequences for a consensus involved in branch formation. Although the data set for this is much smaller than that for splice sites, the consensus sequence WCTAATY was identified, where branch formation occurs at the second A. A good match for this consensus is found within the FKBP12 intron and is indicated in figure 4.9.

By extending the 5' end of the 216/311 cDNA sequence with the genomic sequence from pB201Y we get the combined sequence shown in figure 4.11. This sequence has an ORF of 108 aa, giving it the extra three amino acids that were presumed missing from the original 216/311 protein sequence. The translation is shown under the DNA sequence. An in frame stop codon is found 27 bp upstream from the translation start codon indicating that the ORF does not extend any further upstream. As mentioned earlier, the protein sequence has close homology to a family of 12kDa proteins known as FK binding proteins, for which no *Drosophila* homologue had been identified. Therefore, the sequence in figure 4.11 was submitted to Genbank as the *Drosophila* homologue of FKBP12, and can be viewed under accession number Z49079. Interestingly, several months later the same gene was identified, cloned, and

CCGCTAGCTGTCATTCCGGCGACGCCGACGAACTGAAATCAACGGTCTTATTTTCCGCTCGGTAATTTGCCAGAAAAAATAATCGAA ATG GGA GTA  
M G V

100 *MunI* 150  
CAA GTA GTT CCA ATT GCT CCT GGT GAT G|GC AGC ACC TAT CCC AAG AAT GGC CAA AAG GTC ACG GTC CAC TAC ACC  
Q V V P I A P G D G S T Y P K N G Q K V T V H Y T

200  
GGC ACC CTG GAC GAT GGC ACC AAG TTC GAT TCG CGC GAC CGC AAC AAG CCA TTC AAG TTC ACC ATC GGC AAG  
G T L D D G T K F D S S R D R N K P F K F T I G K

350 *BglI* 300  
GGC GAG GTC ATC CGT GGC TGG GAT GAG GGA GTT GCC CAG TTG AGC GTC GGC CAG AGC GCC AAG CTG ATT TGC TCG  
G E V I R G W D E G V A Q L S V G Q S A K L I C S

400 *BglI* 450  
CCG GAC TAT GCC TAC GGT AGC CGT GGC CAC CCC GGC GTC ATT CCG CCC AAC TCC ACC CTC ACC TTC GAC GTC GAG  
P D Y A Y G S R G H P G V I P P N S T L T F D V E

500 550  
CTG CTC AAG GTC GAA TAG GCGCACAGGATGCCCAATGTGTATACCCCAACCAATCGCAGGGGTTGCGGACCCGGCACC GGCGACGCGG  
L L K V E \*

650  
AAGAACATATAATCGGAACCAGAACATCCAGCAGCATTTGCCAACCCAAAATAATAATGATAAATTCATACACAGCTCTAAAACCTAATAATTAAAC  
CGCAAAAGGGAACGCATCTTCTACCAATACAAA**TAA**CAATTTATTC AAGTCAAAAAA

Figure 4.11. *Drosophila* FKBP12 gene DNA sequence and translation. Bases 1-96 are genomic sequence from pB201Y. Bases 97-646 are cDNA sequence from clones 216 and 311. Upstream in frame stop codon is underlined. Polyadenylation signal is highlighted in bold. Position of intron is indicated by vertical line in the DNA sequence. Positions of *MunI* and *BglI* restriction sites used in orienting the gene are indicated.

sequenced by a different group (Wang *et al.*, 1996), and their sequence can be viewed under accession number U41441. The Wang *et al.* DNA sequence is identical to mine, except that their published sequence has no 5' untranslated region, and has an additional C at position 488 (position 401 of their sequence) in the 3' UTR. And, obviously, the Wang *et al.* protein sequence is identical to mine.

The FKBP family of proteins has a diverse array of members of varying molecular weights from 12kDa up to 52kDa (Galat, 1993). The higher MW members often have a domain similar in sequence to FKBP12, fused to other protein domains of various functionality. It is clear that the sequence described here is a member of the 12kDa subfamily of FKBP, and an alignment between its protein sequence and other sequences from selected members of the FKBP12 subfamily of proteins is shown in figure 4.12. The *Drosophila* sequence is 77% identical to the mouse sequence (with a very similar level of identity to other mammalian sequences), 62% identical to the yeast sequence, and 47% identical to the *Neurospora* sequence. Thus the FKBP12 family of proteins appear to be highly conserved over a wide range of species. Also included in figure 4.12 is the sequence of the FKBP domain of a 39 kDa protein, that is the only other FKBP that has been identified in *Drosophila*. The *Drosophila* FKBP12 sequence is only 50% identical to this sequence (i.e. significantly less than its identity to mammalian FKBP12 sequences), suggesting that the family of FKBP began diverging before the evolutionary split between insects and mammals

#### 4.2.3 Genomic Southern

To determine if the FKBP12 gene is a single copy gene in *Drosophila* a genomic Southern was performed. Genomic DNA from wild-type Oregon-R flies was digested separately with restriction enzymes *SstI* and *XhoI*. Digested DNA was separated by gel electrophoresis, blotted onto nylon filters and probed with the cDNA from pB216 labelled with <sup>32</sup>P. The result is shown in figure 4.13. As can be seen, a single *XhoI* band of 2.4kb (which is the size expected from the pB201Y restriction map) and a



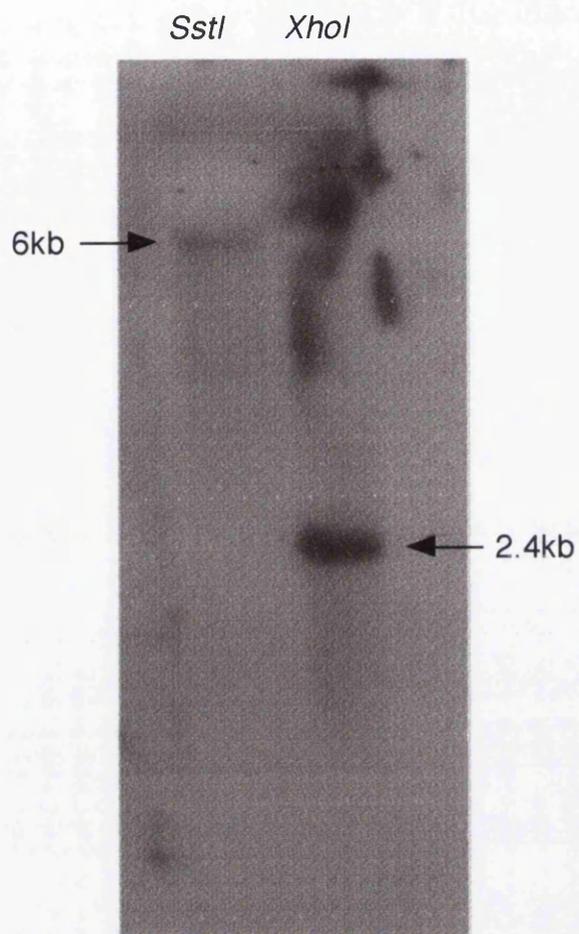


Figure 4.13. Southern blot of wild-type genomic DNA, digested with *SstI* or *XhoI*, and probed with cDNA 311.

single *SstI* band of 6kb hybridise to the probe, suggesting that only a single copy of the FKBP12 gene exists in the *Drosophila* genome.

#### 4.2 4 Orientation of the FKBP12 gene

Does the FKBP12 gene have the same orientation as the P[GAL4] element, i.e. is the gene transcribed from the same strand, and therefore in the same 5' to 3' direction, as the *Gal4* gene of the adjacent enhancer-trap? Whether 201Y enhancer-trap expression reflects that of FKBP12 expression or not, such information is useful in adding to our understanding of which enhancers are most likely to interact with enhancer-traps. Therefore, to answer this question two unique restriction sites, *MunI* and *BglI*, were identified in the FKBP cDNA sequence and are indicated in figure 4.11. If these same restriction sites could be mapped to the plasmid pB201Y this would indicate the orientation of the FKBP12 gene relative to the enhancer-trap. As discussed earlier, reverse Northern analysis had limited the transcriptional unit to lie between the two *XhoI* sites at 5.30 and 7.70 of pB201Y, and further suggested that it straddles the *EcoRI* site at 6.60. Therefore, as an initial step, the following single and double restriction digests were performed on pB201Y: *EcoRI*, *MunI*, *BglI*, *MunI* and *EcoRI*, and *BglI* and *EcoRI*. Restriction fragments were separated by gel electrophoresis and the result is shown in figure 4.14. As can be seen this result indicates that pB201Y contains four *MunI* sites and five *BglI* sites. As this result was more complex than anticipated, the gel was Southern blotted onto nylon filters and probed with the 2.4kb *XhoI* fragment of pB201Y labelled with <sup>32</sup>P, to help indicate which restriction fragments were the pertinent ones. The result of this Southern is also shown in figure 4.14. Two *MunI* fragments, of 2.6kb and 1.3kb, hybridise to the probe, as do two *BglI* fragments of 2.8kb and 1.1kb, suggesting that there is only one restriction site for each enzyme within the 2.4kb *XhoI* fragment of pB201Y. The *MunI*/*EcoRI* double digest yields three fragments of 2.2kb, 1.3kb and 0.4kb which hybridise to the probe. Thus the 2.6kb *MunI* fragment becomes two fragments of 2.2kb and 0.4kb upon

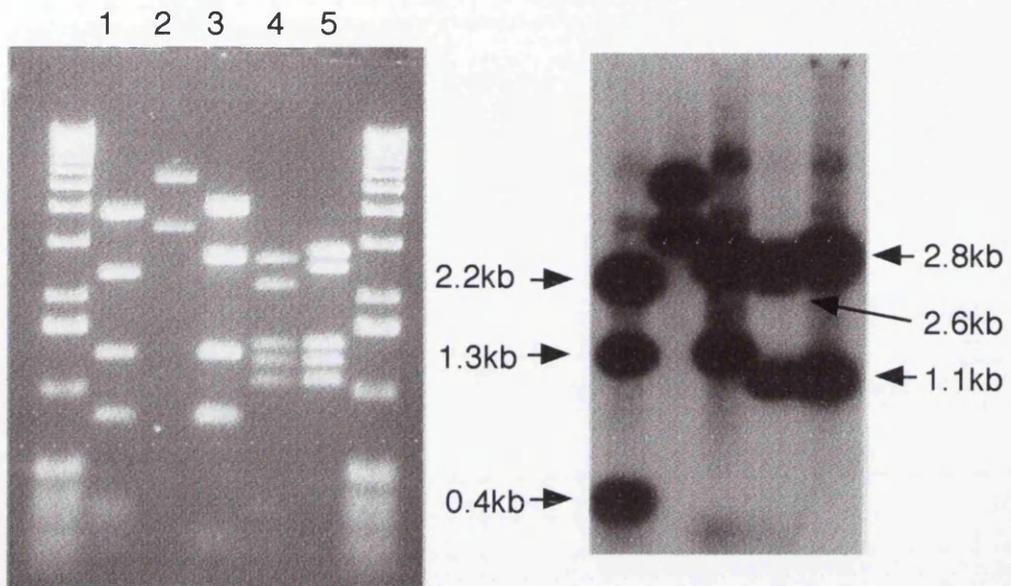


Figure 4.14. Restriction digest of pB201Y. Lane numbers correspond to the following digests: 1=*MunI* + *EcoRI*, 2=*EcoRI*, 3=*MunI*, 4=*BglI* + *EcoRI*, 5=*BglI*. Fragments sized using BRL 1kb ladder. Gel was Southern blotted and probed with the 2.4kb *XhoI* fragment of pB201Y. Result is shown on the right. Sizes and positions of bands discussed in the text are indicated.

double digestion with *MunI* and *EcoRI*, suggesting the *MunI* site is 0.4kb from the *EcoRI* site at 6.60. The *BglII/EcoRI* double digest yields two fragments of 2.75kb and 1.1kb which hybridise to the probe, suggesting that the *BglII* site is very close to the *EcoRI* site (note that no hybridisation to a third *BglII/EcoRI* fragment is evident presumably due to the small size of this fragment). Therefore the critical question is whether the *MunI* site is at 7.00 or 6.20 on the restriction map shown in figure 4.4. If it is at 7.00 one would expect to see *MunI/XhoI* fragments of 1.7kb and 0.7kb, while if it is at 6.2, one would expect *MunI/XhoI* fragments of 1.5kb and 0.9kb, upon digestion of pB201Y with *MunI* and *XhoI*. The result of such a digest is shown in figure 4.15. There are clearly *MunI/XhoI* fragments of 1.7kb and 0.7kb with no fragments evident at 1.5kb or 0.9kb, therefore placing the *MunI* site at 7.00. With this information we can orient the FKBP12 relative to the P[GAL4] element. This is shown in figure 4.16, where it can be seen that the FKBP12 gene has the opposite 5' to 3' orientation relative to the P[GAL4] element.

An examination of the 216/311 cDNA sequence indicates the lack of an *EcoRI* restriction site, that should be present based on the pB201Y restriction map. The intron sequence discussed earlier also does not contain an *EcoRI* site. However, it is possible that the *EcoRI* site resides within this intron as the sequence was obtained from a single sequencing run, and potential sequencing errors may be obscuring the presence of the *EcoRI* site. And indeed there are two points where the intronic sequence is only one base different from the GAATTC restriction site for *EcoRI*. Of course the simplest explanation is that the *EcoRI* site resides within a yet unidentified intron.

#### **4.2.5 Expression of the FKBP12 gene**

To look at the expression of FKBP12 a developmental Northern was performed using total RNA prepared from unstaged embryos, larvae, and pupae, and adult heads and bodies. Equal amounts of RNA (approximately 10µg) were loaded onto a denaturing gel, separated by electrophoresis, and blotted onto nitrocellulose filters. Filters were

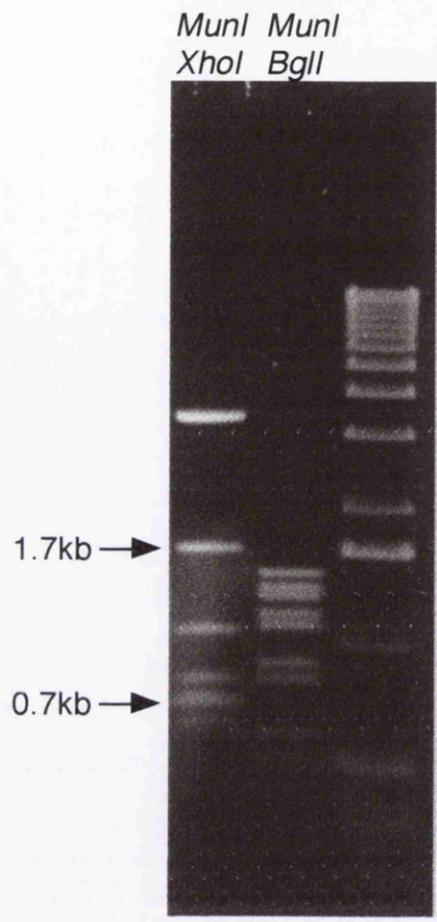


Figure 4.15. Restriction digests of pB201Y. Enzymes used are indicated above the lanes. Fragments sized using BRL 1kb ladder. 1.7kb and 0.7kb *MunI/XhoI* fragments are indicated.



probed with the cDNA fragment from pB216 labelled with  $^{32}\text{P}$ , and exposed to X-ray film for 24hrs at  $-70^{\circ}\text{C}$ . The result is shown in figure 4.17. A single band of approximately 700bp is apparent for all lanes, indicating that a single FKBP12 transcript is expressed at all developmental stages. The size of the transcript identified is in good agreement with that suggested by the cDNA sequence (a total of 610 bp of coding sequence and 3' UTR). Expression appears to be relatively uniform throughout development with no particular stage having elevated expression. Also expression does not appear to be particularly elevated in the head relative to the body, which is in agreement with the reverse Northern data discussed earlier, further suggesting that *Drosophila* FKBP12 is not concentrated in neuronal tissue.

To look more closely at FKBP12 expression, a polyclonal antibody raised to the polypeptide DVELLKLE was acquired (gift from G. Wiederrecht, Merck Research Laboratories). This polypeptide represents the final eight amino acids of the carboxy terminus of the bovine FKBP12. The antibody had been used successfully to identify FKBP12s in Western blots of calf brain extracts (Sewell *et al.*, 1994). However, when used immunohistochemically to analyse the expression of FKBP12 in whole brains from *Drosophila* adults, no specific reactivity could be detected. This may be because the antibody does not cross-react with the *Drosophila* protein, as the peptide sequence against which it was raised differs slightly from the corresponding *Drosophila* sequence, which is DVELLKVE. However, the change of a leucine to a valine at the penultimate position is relatively minor, both having similar nonpolar side chains, and therefore may not be sufficient to account for the lack of reactivity observed. Alternatively, it may be that the particular peptide sequence recognised by the antibody is not accessible in the intact protein or is in a conformation that makes it unreactive. And indeed to my knowledge this antibody has not been used successfully to study the distribution of intact FKBP12s in whole tissues.

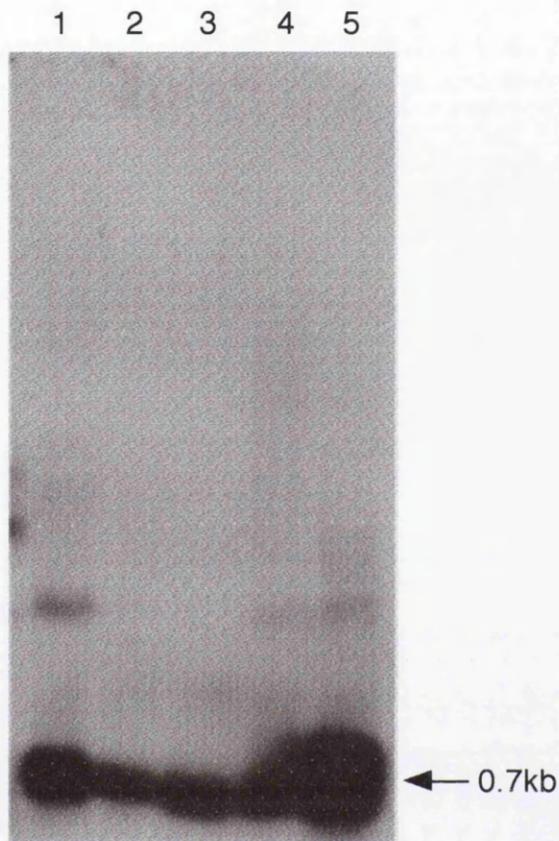


Figure 4.17. Developmental Northern. Sources of RNA as follows: 1 = Embryos, 2 = Pupae, 3 = Larvae, 4 = Adult heads, 5 = Adult bodies. Probed with cDNA 216.

## 4.3 Discussion

### 4.3.1 FK Binding Proteins

FK binding proteins (FKBPs) are a family of peptidylprolyl isomerases (PPIases) that catalyse the interconversion of *cis* and *trans* isomers of peptidyl-prolyl amide bonds (EC 5.2.1.8). The isomerase activity of FKBPs has been shown to accelerate protein folding *in vitro* (Fischer and Schmid, 1990). The cyclophilins, a separate family of proteins of unrelated sequence, also possess PPIase activity, and are grouped with the FKBPs under the general heading of immunophilins (reviewed in Galat, 1993). Interest was originally focussed on FKBPs as the intracellular receptors for the immunosuppressant macrolides FK506 (hence the name FK binding protein) and rapamycin. Both FK506 and rapamycin inhibit FKBP catalytic activity, by binding with high affinity ( $K_{ds}$  of 0.4nM and 0.2nM respectively), to a deep hydrophobic pocket within the protein that also contains the isomerase active site (Michnick *et al.*, 1991). Remarkably, the other family of PPIases, the cyclophilins, are receptors for the immunosuppressant cyclosporin A.

The first FKBP identified was a 12kDa version (known as FKBP12) isolated independently from human spleen and T-lymphoma cells (Harding *et al.*, 1989; Siekierka *et al.*, 1989). Since then a large number of similar and related proteins have been isolated from a variety of eukaryotic and prokaryotic species, although only members of the FKBP12 subgroup have been firmly implicated in immunosuppression. Many of the larger FKBPs, apart from containing the core FKBP domain, contain other domains held in common with closely related FKBPs. Furthermore, some of these other domains are homologous to other proteins. For instance, the rabbit 52kDa FK binding protein has a domain homologous to a mitochondrial outer membrane protein, plus another domain homologous to a yeast stress inducible protein (Lebeau *et al.*, 1992).

The macrolide FK506 is known to suppress the human immune system by preventing T-cell activation (Crabtree, 1989; Isakov *et al.*, 1987). In particular, FK506 blocks a step required for the activation of transcription factors which promote expression of lymphokine genes (Tocci *et al.*, 1989). The transcriptional activation, and (to a lesser degree) DNA-binding properties, of the transcription factors NF-AT, AP-3, and (partially) NF- $\kappa$ B are all potently inhibited by FK506, but not rapamycin (Mattila *et al.*, 1990).

Recently the first *Drosophila* gene encoding a 39kDa member of the FKBP family of proteins was cloned (Theopold *et al.*, 1995). This FKBP is expressed predominantly in the early (0-12h) embryo, in the ovaries of female adults, and also in the *mbn-2* haemocyte line. This cell line was derived from a blood tumour mutant, and has been used as a model of the *Drosophila* immune response, as in the presence of microbial substances the cells respond by activating genes for antimicrobial peptides. However, FK506 did not interfere with the immune response of this cell line, suggesting that FKBP39, and presumably FKBP12, do not play a role in the activation of the *Drosophila* antibacterial response.

As mentioned in the results section the *Drosophila* FKBP12 was also cloned and sequenced by Wang *et al.*, (1996). It had previously been shown that TGF $\beta$  type I serine/threonine kinase receptors interact specifically with FKBP12, in a manner suggesting that TGF $\beta$  type I receptors are natural ligands for the FK506 binding site (Wang *et al.*, 1994). In this further study they identified, in a yeast two-hybrid screen, a *Drosophila* cDNA from an imaginal disc library, as an interactor with both the *Drosophila* TGF $\beta$  type I receptors Thickvein and Sax. This cDNA encodes a FKBP12 and its close similarity to the FKBP12 that I identified was discussed in the results section. Furthermore, they showed that human FKBP12 binding to TGF $\beta$  type I receptors inhibits TGF $\beta$  signalling pathways by recruiting another cytoplasmic protein, possibly calcineurin.

It is known that the cytoplasmic concentration of FKBP may approach 5nM in T-cells, whereas inhibition of T-cell proliferation occurs at subnanomolar

concentrations of FK506 (Dumont *et al.*, 1990). Therefore, it is unlikely that the inhibition of FKBP isomerase activity by FK506 mediates the suppression of T-cell activation, as only a small fraction of the enzyme is inhibited at effective drug concentrations. Adding weight to this conclusion, is the observation that compounds related to FK506 (e.g. L-685, 818 and FK506BD) inhibit FKBP12 isomerase activity, but are not themselves immunosuppressive (Bierer *et al.*, 1990). In yeast, deletion of the FKBP12 gene results in rapamycin resistant strains that are returned to rapamycin sensitivity following transfection with human FKBP12 (Koltin *et al.*, 1991). Thus it would appear that binding of immunosuppressants to FKBP12 results in a gain, rather than loss, of function.

Using a glutathione S-transferase-FKBP12 fusion protein, immobilised on a glutathione-sepharose column, proteins that bind to the FK506/FKBP12 complex in calf brain, spleen, and thymus extracts have been examined (Liu *et al.*, 1991). Proteins of  $M_r$  61000, 57000, 17000 and 15000 were isolated from all three tissue types. None of these proteins were isolated in the absence of FK506 or  $Ca^{2+}$ . Subsequent biochemical, immunochemical, and sequence analysis revealed the 17kDa protein to be calmodulin, the 61kDa protein to be calcineurin A, and the 15kDa protein to be calcineurin B (the 57kDa protein was a proteolytic product of calcineurin A). Interestingly, similar experiments with a complex of cyclophilin and the immunosuppressant cyclosporin A isolated identical proteins.

#### **4.3.2 FKBP12 and Calcineurin**

Calcineurin, also known as protein phosphatase 2B (PP2B), is a  $Ca^{2+}$ , calmodulin dependent serine/threonine protein phosphatase (Stewart *et al.*, 1982). It is part of a larger family of phosphatases that includes protein phosphatase 1 (PP1) and protein phosphatases 2A and 2C. The phosphatase activity of calcineurin towards a phosphopeptide substrate, is potently inhibited by the FK506/FKBP12 complex, but not by the rapamycin/FKBP12 complex nor by FKBP12 alone (Liu *et al.*, 1991).

Interestingly calcineurin dephosphorylation of p-nitrophenyl phosphate is enhanced by FK506/FKBP12, suggesting that modulation of calcineurin activity is achieved by binding to a site adjacent to the active site.

Calcineurin is a heterodimer consisting of a large ( $M_r$  58000-59000) catalytic A subunit, and a small ( $M_r$  19000)  $Ca^{2+}$  binding regulatory B subunit. Binding of  $Ca^{2+}$  to calcineurin B promotes a small basal activity, whereas  $Ca^{2+}$  binding to calmodulin facilitates calmodulin's interaction with calcineurin A and results in 10-fold activation of the enzyme (Klee *et al.*, 1987). The catalytic activity of the enzyme is confined to the amino-terminal two-thirds of calcineurin A. The carboxyl-terminal third contains a regulatory domain, consisting of calmodulin binding and calcineurin B binding subdomains, and an autoinhibitory subdomain. The inactive conformation of the enzyme, maintained by the autoinhibitory domain, is relieved by  $Ca^{2+}$ -dependent binding of calmodulin (Hubbard and Klee, 1989).

*Drosophila* genes encoding the calcineurin A and B subunits have been cloned. Calcineurin B is the product of a single gene, and the *Drosophila* protein is 88% identical to the human form. The highly conserved nature of calcineurin B suggests that few residues can be altered without affecting function. In common with other species, there is more than one isoform of *Drosophila* calcineurin A, and to date two separate calcineurin A genes have been characterised (Guerini *et al.*, 1992; Brown *et al.*, 1994.) Furthermore, additional isoforms have been suggested by PCR analysis (Chen *et al.*, 1992).

Following the discovery that the Fk506/FKBP12 complex inhibits calcineurin phosphatase activity, a mechanism whereby FK506 prevents T-cell activation was rapidly worked out. Following antigen recognition by the T-cell receptor, a kinase cascade causes the generation of inositol triphosphate and diacyl glycerol. The former molecule causes a release of  $Ca^{2+}$  from intracellular stores with the consequent activation of calcineurin. Translocation to the nucleus of the transcription factor NF-AT, following dephosphorylation by calcineurin, results in the activation of early T-cell gene products. The FK506/FKBP12 complex by inhibiting calcineurin prevents

dephosphorylation of NF-AT and it presumably remains in the cytosol (Schreiber and Crabtree, 1992).

Although much research has been focussed on the role of FKBP12 and calcineurin in immunosuppression, calcineurin has a higher expression in the brain (comprising 1% of total protein in some regions) than in tissues of the immune system (Klee and Cohen, 1988). The distribution of FKBP12 in rat brain has been determined by measuring [<sup>3</sup>H]FK506 binding to sections and homogenised fractions. This revealed that FKBP12 is highly expressed in the cortex, striatum, and hippocampus at levels more than ten-fold greater than in spleen or thymus, and also in other brain structures but at much lower levels (Steiner *et al.*, 1992). Subsequent studies using *in situ* hybridisation to brain sections, with antisense oligonucleotide probes complementary to a FKBP12 cDNA, confirmed these results. Furthermore, brain distributions of calcineurin, determined by immunohistochemistry, were shown to closely mirror that of FKBP12 (Dawson *et al.*, 1994). These observations fit with the notion that FKBP12 exerts its major biological action through inhibition of calcineurin activity. The very high levels, and discrete localisations, in the brain of FKBP12 and calcineurin indicate some key regulatory function for these proteins. As FKBP12 inhibits calcineurin activity only when bound to FK506, the existence of endogenous ligands is strongly implied (although to date none have been identified). Although it does not appear that *Drosophila* FKBP12 has elevated expression in the head or in neuronal tissue, it no doubt has a function within the *Drosophila* CNS.

There have been a number of insights into the possible functional roles of calcineurin and/or FKBP12 in neurons. Long-term potentiation (LTP) was blocked in hippocampal CA1 pyramidal cells in the presence of a strong inhibitor of calcineurin, but not in the presence of a weak inhibitor (Wang and Stelzer, 1994). Application of FK506 to hippocampal slices post-synaptically also blocks long-term depression (LTD) induced by prolonged electrical stimulation (Mulkey, *et al.*, 1994). Other calcineurin inhibitors elicit a similar effect, and dephosphorylation of inhibitor-1 by calcineurin is thought to be the crucial step. Phosphorylated inhibitor-1 inhibits PP1 activity, thus

FK506 prevents LTD by ultimately preventing PP1 dephosphorylating necessary substrates. Another inhibitor of PP1 - dopamine and cAMP-regulated phosphoprotein of MW 32kDa (DARPP-32) - is dephosphorylated and inactivated by calcineurin following NMDA receptor mediated  $\text{Ca}^{2+}$  influx in rat striatal neurons (Halpain *et al.*, 1990). Calcineurin inactivation of DARRP-32 leads, via activation of PP1, to decreased phosphorylation and activation of  $\text{Na}^+, \text{K}^+$ -ATPase (Aperia *et al.*, 1992). Application of FK506 inhibited activation of this neuronally ubiquitous cation pump. FK506 reduces glutamate neurotoxicity in primary cerebral cortical neuronal cultures by enhancing phosphorylation of nitric oxide synthase (NOS) (Dawson *et al.*, 1993). The proposed mechanism also involves an influx of  $\text{Ca}^{2+}$  activating calcineurin, following glutamate activation of NMDA receptors, and the subsequent dephosphorylation and activation of NOS. FK506/FKBP12, by inhibiting calcineurin, maintains NOS in its inactive phosphorylated form, thus reducing the production of nitric oxide (which is neurotoxic at high concentrations, figure 4.18).

FKBP12 levels are elevated in peripheral neurons undergoing regeneration following physical insult (Lyons *et al.*, 1995). Furthermore, there is an associated increase in levels of growth associated protein-43 (GAP43), a known substrate of calcineurin (Steiner *et al.*, 1992), although calcineurin itself is not elevated in these regenerating neurons. Very low levels (1pM) of FK506 promotes neurite outgrowth of rat dorsal root ganglion cells and pheochromocytoma cells (Lyons *et al.*, 1994). The effect of FK506 is mediated by FKBP, and is dependent on the increased sensitivity of the cells to nerve growth factor. The possibility that FK506/FKBP inhibition of calcineurin is responsible for neurite outgrowth, is suggested by the fact that numerous proteins are phosphorylated in neurite extensions (Levi *et al.*, 1988). Inhibitors of calcineurin prevent axonal elongation in developing neurons derived from embryonic rat cerebellum (Ferreira *et al.*, 1993). Furthermore, inhibition of calcineurin led to a *decrease* in the phosphorylation of tau, a microtubule-associated protein thought to play a key role in neurite elongation. Inhibition of calcineurin by an FK506 analogue reduced stimulation-induced dephosphorylation of another microtubule-associated

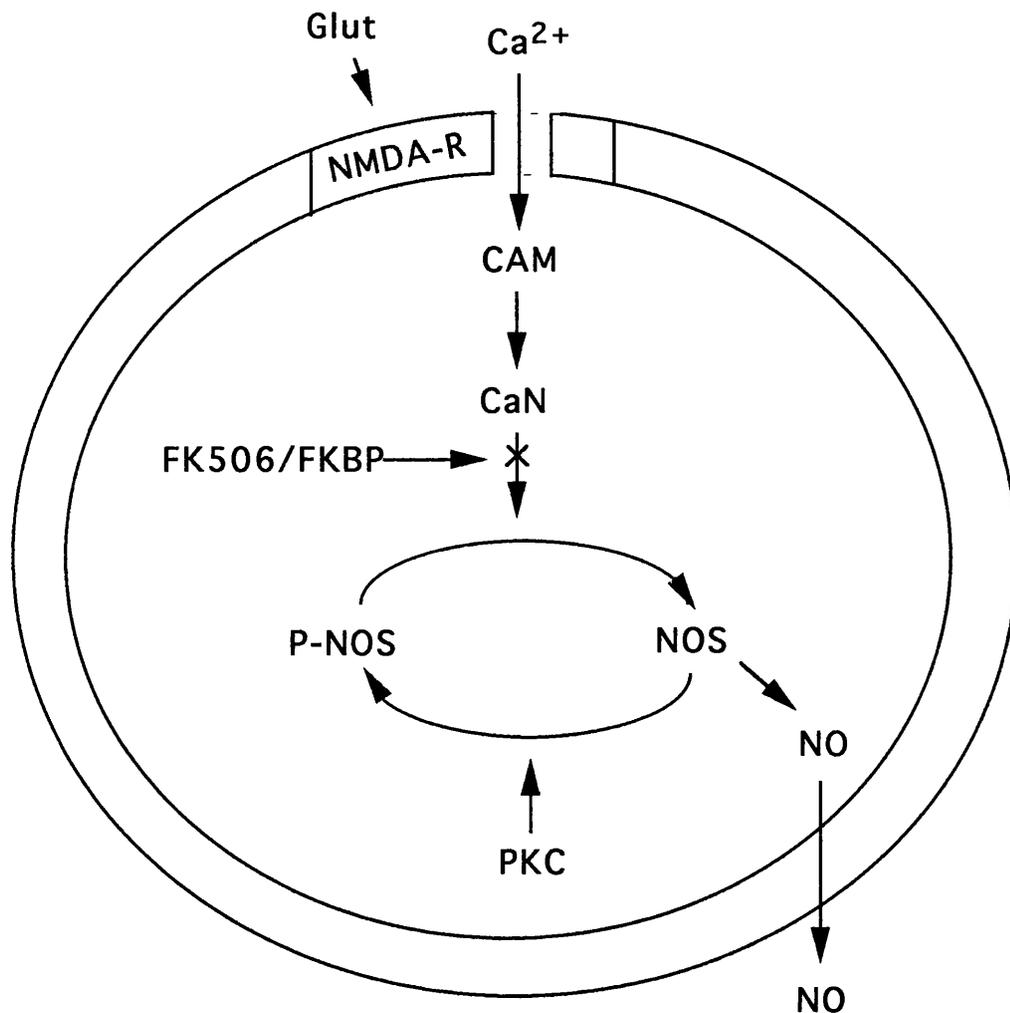


Figure 4.18. Regulation of neuronal NOS activity by calcineurin (CaN). Ca<sup>2+</sup> entry, via NMDA receptors, activates CaN, which dephosphorylates and activates NOS leading to NO production. The FK506/FKBP complex inhibits CaN and prevents dephosphorylation of NOS leading to reduced NOS activity and lowered NO production. (Adapted from Dawson *et al.*, 1993)

protein, the motor-protein dynamin (Nichols *et al.*, 1994). The same study found that calcineurin inhibition led to an enhancement of stimulation-induced glutamate release from rat brain nerve terminals. The microtubule-associated protein, MAP1B, is a substrate of calcineurin *in vitro* (Ulloa *et al.*, 1993), as are the protein kinase C substrates, neurogranin, neuromodulin, and MARCKS (Seki *et al.*, 1995)

Many substrates for calcineurin are phosphorylated by cAMP-dependent protein kinase (PKA), suggesting a mechanism whereby  $\text{Ca}^{2+}$ , by activating calcineurin, antagonises the actions of cAMP. In addition, dephosphorylation of inhibitor-1 and DARPP-32 would lead to activation of PP1 and dephosphorylation of yet other PKA substrates (figure 4.19). Furthermore, the activity of PKA is directly modulated by calcineurin, as the  $R_{II}$  regulatory subunit of PKA is itself a substrate of calcineurin (Blumenthal *et al.*, 1986). Thus, calcineurin dephosphorylation of the autophosphorylated  $R_{II}$  subunit increases the rate of reassociation with the catalytic subunit, with the consequent inactivation of PKA. The regulation of the phosphorylation state of common substrates, by calcineurin and PKA, is emphasised by their association with a common anchoring protein, AKAP79. This protein localises PKA to post-synaptic densities, and also binds calcineurin in cultured hippocampal neurons (Coghlan *et al.*, 1995).

#### **4.3.3 FKBP's and Ryanodine Receptors**

FKBP12 copurifies with the ryanodine receptor (RyR, named after its specific agonist ryanodine), a large  $\text{Ca}^{2+}$  channel that controls release of  $\text{Ca}^{2+}$  from  $\text{Ca}^{2+}$ -storing organelles like the endoplasmic reticulum and muscle sarcoplasmic reticulum (Jayaraman *et al.*, 1992). RyRs are structurally and functionally related to inositol 1,4,5-triphosphate receptors ( $\text{IP}_3\text{Rs}$ ), and are formed from 550-565kDa monomers which aggregate into functional tetrameric complexes. Three isoforms of RyRs are known, skeletal muscle (or type I), cardiac muscle (or type II), and brain (or type III). The type I isoform is sensitive to cellular depolarisation, whilst the type II

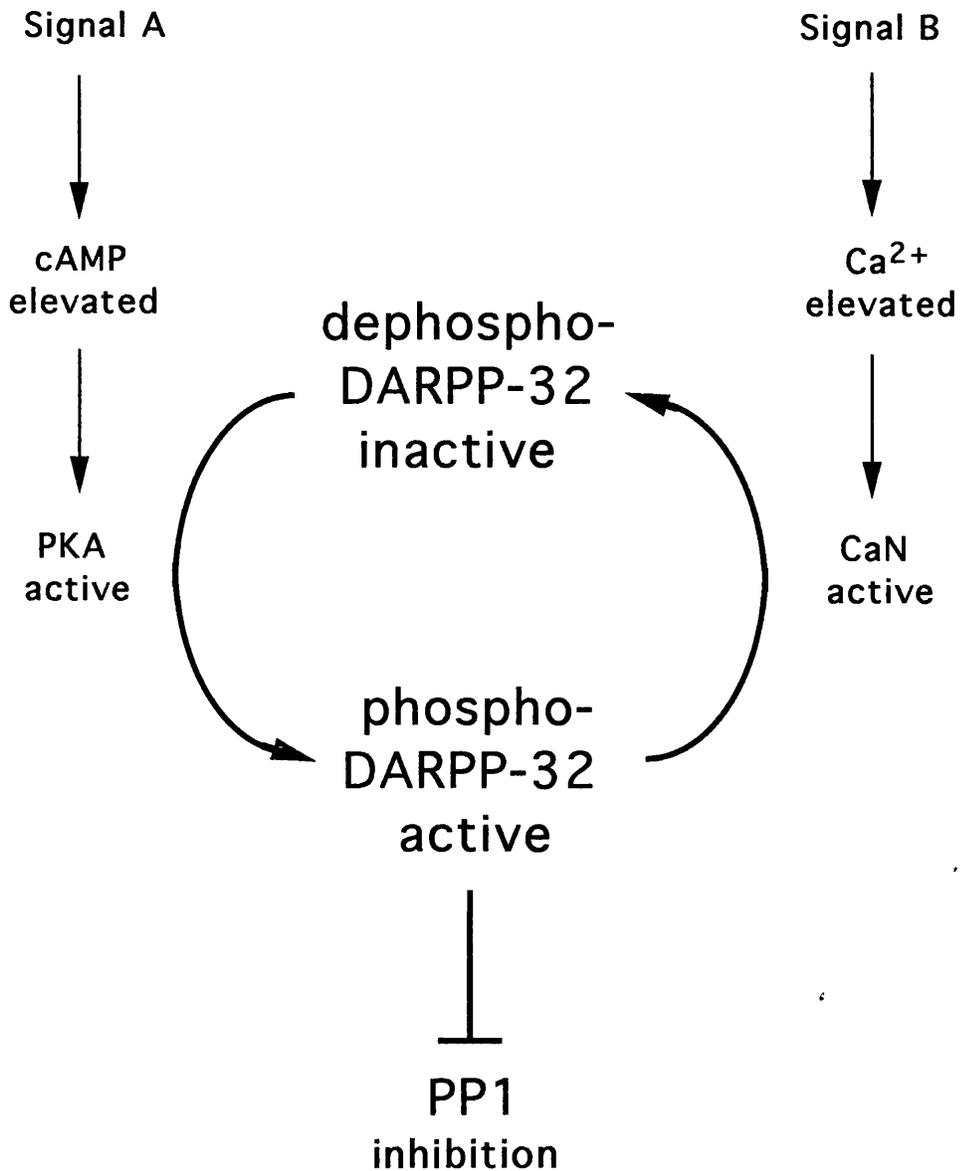


Figure 4.19. Antagonistic activities of calcineurin (CaN) and PKA. Signal A leads to an increase in cAMP and activation of PKA, leading to phosphorylation and activation of DARPP-32. Signal B leads to an increase in Ca<sup>2+</sup> and activation of CaN, causing dephosphorylation and inactivation of DARPP-32. This removes the inhibition of PP1 leading to the dephosphorylation of PKA substrates. (Adapted from Cohen, 1992).

and type III isoforms are activated by  $\text{Ca}^{2+}$ , allowing  $\text{Ca}^{2+}$ -induced  $\text{Ca}^{2+}$  release (CICR) (reviewed in Meissner, 1994). Although the brain possesses its own specific isoform of RyR, all three isoforms are present in brain tissue, with type II predominating and types I and III restricted to specific regions (Furuichi *et al.*, 1994).

Each RyR monomer has associated with it one molecule of FKBP12 (thus, there are four molecules of FKBP12 per functional  $\text{Ca}^{2+}$  channel) (Timmerman *et al.*, 1993). By expressing a RyR cDNA in insect Sf9 cells the effect of FKBP12 on RyR function has been studied (Brilliantes *et al.*, 1994). In the absence of FKBP12 the caffeine-activated RyR displayed increases in the probability of opening, and multiple subconductance states making the channel "leaky". In the presence of FKBP12 the activated RyR displayed a lower probability of opening, but when open, full conductance states are stably formed. Thus, FKBP12 appears to stabilise the channel, making it harder to open but more stable once it is open, possibly by improving cooperativity among the four RyR subunits. FK506 or rapamycin reverses this stabilising effect by dissociating the FKBP from the RyR (Timmerman *et al.*, 1993). The possibility that the isomerase activity of FKBP is important in their stabilisation of the RyR immediately suggests itself. However, substitution of wild-type FKBP12 on the RyR with mutant FKBP12 deficient in PPIase activity, did not alter the  $\text{Ca}^{2+}$  conductance of the channel in terminal cisternae vesicles (Timmerman *et al.*, 1995). Thus, the modulation of RyR function by FKBP12 appears to be independent of its PPIase activity.

Recently the association of FKBP12 with the other major intracellular  $\text{Ca}^{2+}$  release channel, the  $\text{IP}_3$  receptor, has been suggested (Cameron *et al.*, 1995). It would appear that FKBP12 has a similar stabilising effect on the  $\text{IP}_3\text{R}$ , as addition of FK506 or rapamycin causes leakiness of the channels gating properties. It is a remarkable finding that both of the major intracellular  $\text{Ca}^{2+}$  release channels, the  $\text{IP}_3\text{R}$  and the RyR, have their function modulated by FKBP12. Calcium is a major cellular second-messenger and intracellular  $\text{Ca}^{2+}$  release channels participate in a vast array of cell signalling pathways including fertilisation, neurotransmitter release, hormonal

activation, T-cell activation, and excitation-contraction coupling in muscle (Berridge, 1993). There is an undoubted role for  $\text{Ca}^{2+}$  in the cellular processes underlying learning and memory, with the following points worthy of note. The posttranslational modification of preexisting proteins by  $\text{Ca}^{2+}$ -dependent enzymes is involved in the induction of synapse-specific changes, such as those that accompany LTP and LTD. The role of CREB in LTM formation is well documented, and its activity is modulated by  $\text{Ca}^{2+}$  dependent kinases, such as CaM kinases II and IV (Dash *et al.*, 1991).

Whilst there is no direct evidence to suggest that FKBP12 has elevated expression in the mushroom bodies of *Drosophila*, it is clear from the previous discussion that it could function in those structures to modulate learning and memory. Calcineurin may potentially dephosphorylate the substrates of cAMP dependent protein kinase (PKA), and the inhibition of calcineurin by FKBP12 could be critical in allowing the physiological changes wrought by PKA during associative learning (see section 1.7). Furthermore, the regulation of calcium release from intracellular stores by FKBP12 would directly impinge on the activity of adenylyl cyclase, an enzyme known to have a crucial role in *Drosophila* learning and memory

#### **4.3.4 FKBP's and Rapamycin**

Although the interaction of the immunosuppressant complex FK506/FKBP12 with its major cellular target, calcineurin, has been well studied, the cellular targets of the immunosuppressant complex rapamycin/FKBP12 are less well understood. The immunosuppressant activity of rapamycin acts via a different mechanism to FK506, and correlates best with a decreased activity of the 70kD S6 kinase which phosphorylates the S6 protein of the small ribosomal subunit (Price *et al.*, 1992). However, the S6 kinase is not thought to be the direct target of the rapamycin/FKBP12 complex. In yeast the lethal effects of rapamycin are eliminated by mutations in two proteins known as TOR1 and TOR2 (Kunz *et al.*, 1993). Mammalian analogues of the TOR proteins have been recently identified. RAFT1, a 289kD protein,

that specifically binds the rapamycin/FKBP12 complex has 40% identity to TOR1 and TOR2 (Sabatini *et al.*, 1994). Both RAFT1 and yeast TOR proteins have considerable sequence similarity to the 110kD subunit of the phosphatidylinositol-3-OH kinase (PI3K). PI3K is known to be associated with stimuli that alter cell growth and division, it seems reasonable to suggest that inhibition of similar RAFT1 activity may mediate rapamycin's immunosuppressant activity.

**Chapter 5**  
**The 201Y 5' Gene**

## 5.1 Introduction

As outlined in Chapter 3, plasmid-rescue on genomic DNA from fly line 201Y yielded only a 3' plasmid. No 5' plasmid could be generated, presumably because of the lack of a suitable *KpnI* restriction site. To identify transcriptional units located 5' to the P[GAL4] insertion point, it was first necessary to acquire genomic clones covering this region.

## 5.2 Identification and analysis of the 201Y 5' gene

### 5.2.1 Isolation and Restriction mapping of Genomic clones

To isolate genomic clones, a genomic library, constructed by partial *Sau3A* restriction digest of *Drosophila* genomic DNA inserted into the bacteriophage vector lambda GEM-11, was screened with the 5.9 kb *BamHI* fragment of pB201Y (figure 4.2). Approximately 50,000 recombinant phage were plated out and duplicate filter lifts probed with the *BamHI* fragment labelled with <sup>32</sup>P. Nine positives were identified and purified down to single plaques. Phage DNA was prepared from each positive, and restriction digested with *SstI*, *XhoI*, and *SstI* and *XhoI*. Restriction fragments were separated by gel electrophoresis, and the result is shown in figure 5.1. As can be seen the nine positive phage, represent three different genomic clones. One clone from each of the three different types was selected for further study, i.e. clones 1.1, 2.1, and 4.2.

To construct restriction maps of these selected genomic clones, they were further subjected to the following single and double restriction digests: *XhoI*, *XhoI* and *EcoRI*, *EcoRI*, *SstI* and *EcoRI*, *SstI*, *XhoI* and *SstI*. Again restriction fragments were separated by gel electrophoresis and the results are shown in figure 5.2. To assist with the construction of restriction maps, and to confirm the genomic clones as representing DNA from the 201Y P[GAL4] locus, the gel was Southern blotted onto nylon filters and probed with the 5.5kb *SstI/BamHI* restriction fragment of pB201Y, labelled with

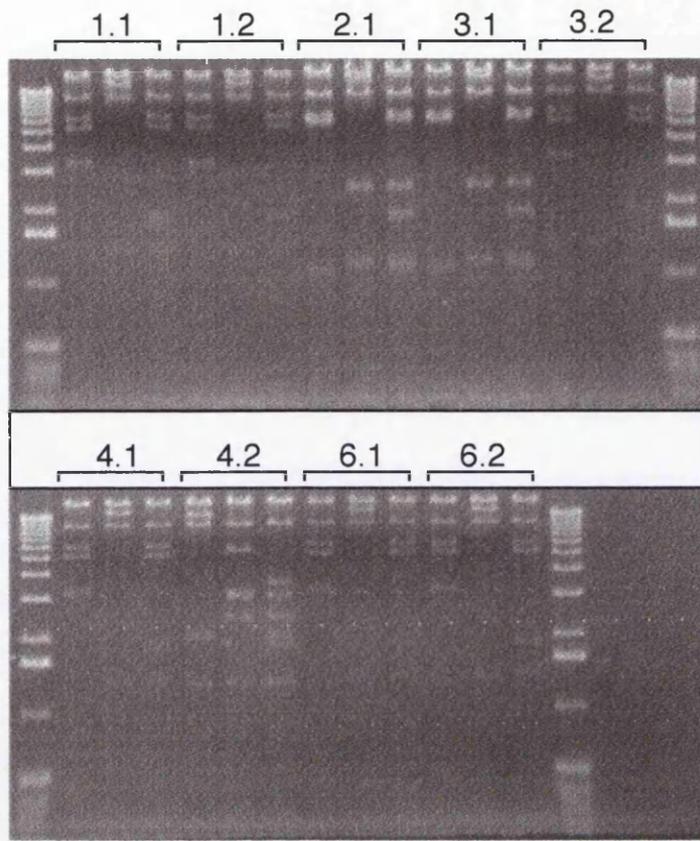


Figure 5.1. Restriction digests of isolated genomic clones. For each clone digests are from left to right: *XhoI*, *SstI*, *XhoI* + *SstI*. Fragments sized using BRL 1kb ladder.

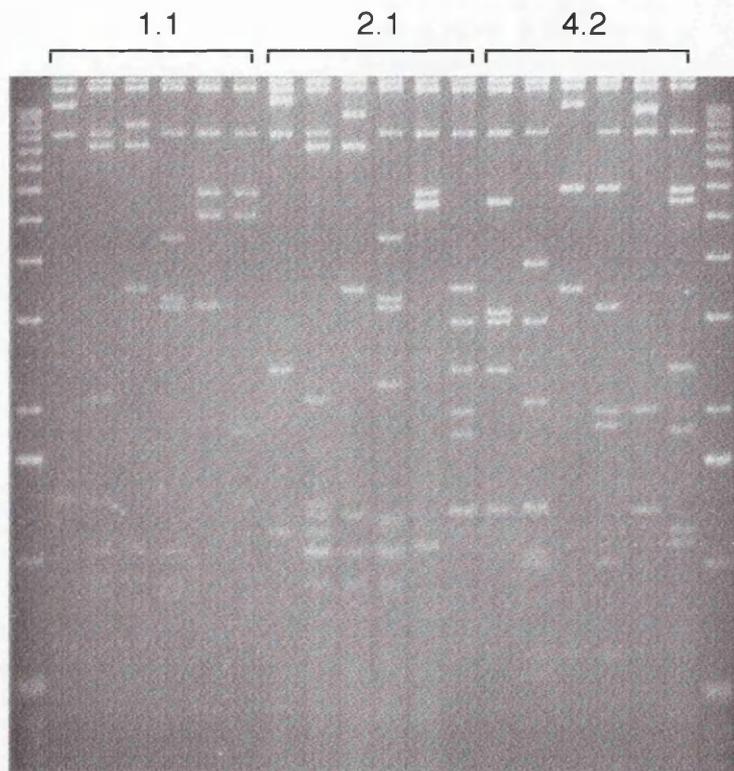


Figure 5.2. Restriction digests of genomic clones 1.1, 2.1, and 4.2. Fragments sized using BRL 1kb ladder. For each clone digests from left to right are: *XhoI*, *XhoI* + *EcoRI*, *EcoRI*, *EcoRI* + *SstI*, *SstI*, *XhoI* + *SstI*. (Note: 2.1 and 4.2 *XhoI/SstI* lanes are transposed.)

<sup>32</sup>P. Probed filters were exposed to X-ray film for 2hrs at -70°C, and results are shown in figure 5.3. From the data contained in figures 5.3 and 5.2 the restriction maps depicted in figure 5.4 can be constructed.

From the restriction maps of figure 5.4 it is proposed that 1.1, 2.1, and 4.2 represent a set of overlapping genomic clones, and that the genomic insert of clone 1.1 is spanned entirely by that of clone 2.1. To verify these regions of overlap each genomic clone was restriction digested with *XhoI* and *SstI*, and restriction fragments loaded onto a gel in triplicate and separated by electrophoresis. Results are shown in figure 5.5. Each set of three digests was Southern blotted onto separate nylon filters and separately probed with either 1.1, 2.1 and 4.2 whole phage DNA labelled with <sup>32</sup>P. Filters were exposed to X-ray film for 1hr at -70°C, and the result is shown in figure 5.6. All of the 1.1 restriction fragments hybridise to the 2.1 probe (but the reverse is not the case) confirming that the 1.1 genomic insert is spanned entirely by that of 2.1. Thus, further analysis was confined to the two genomic clones, 2.1 and 4.2.

### **5.2.2 Reverse Northern analysis of Genomic clones**

To identify transcriptional units within the 5' region of the 201Y P[GAL4] insertion site genomic clone 2.1 was subjected to reverse Northern analysis. Furthermore, to identify additional 3' transcriptional units (other than the FKBP12 gene discussed in chapter 4), genomic clone 4.2 was subject to the same analysis. Phage DNA from clones 2.1 and 4.2 was restriction digested with *XhoI* and *SstI*, and *SstI* and *EcoRI*. Restriction fragments were loaded onto gels in duplicate and separated by electrophoresis (figure 5.7). Each set of digests was Southern blotted onto separate filters, and each filter probed with either <sup>32</sup>P labelled cDNA copied from *Drosophila* head polyA<sup>+</sup> RNA or from *Drosophila* body polyA<sup>+</sup> RNA. Filters were exposed to X-ray film for 19hrs at -70°C. Results are shown in figure 5.8. As can be seen a numbers of bands from both genomic clones hybridise to both probes. Many of these bands correspond to restriction fragments that are known to contain the FKBP12 gene

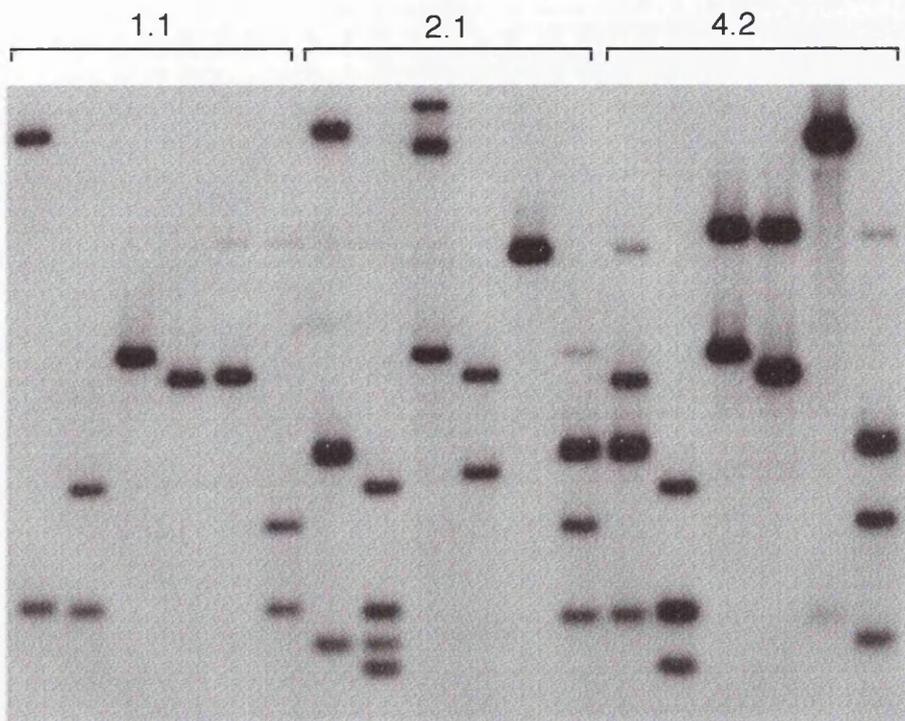


Figure 5.3. Southern blot of gel shown in figure 5.2, probed with the 5.5kb *SstI/XhoI* fragment of pB201Y.

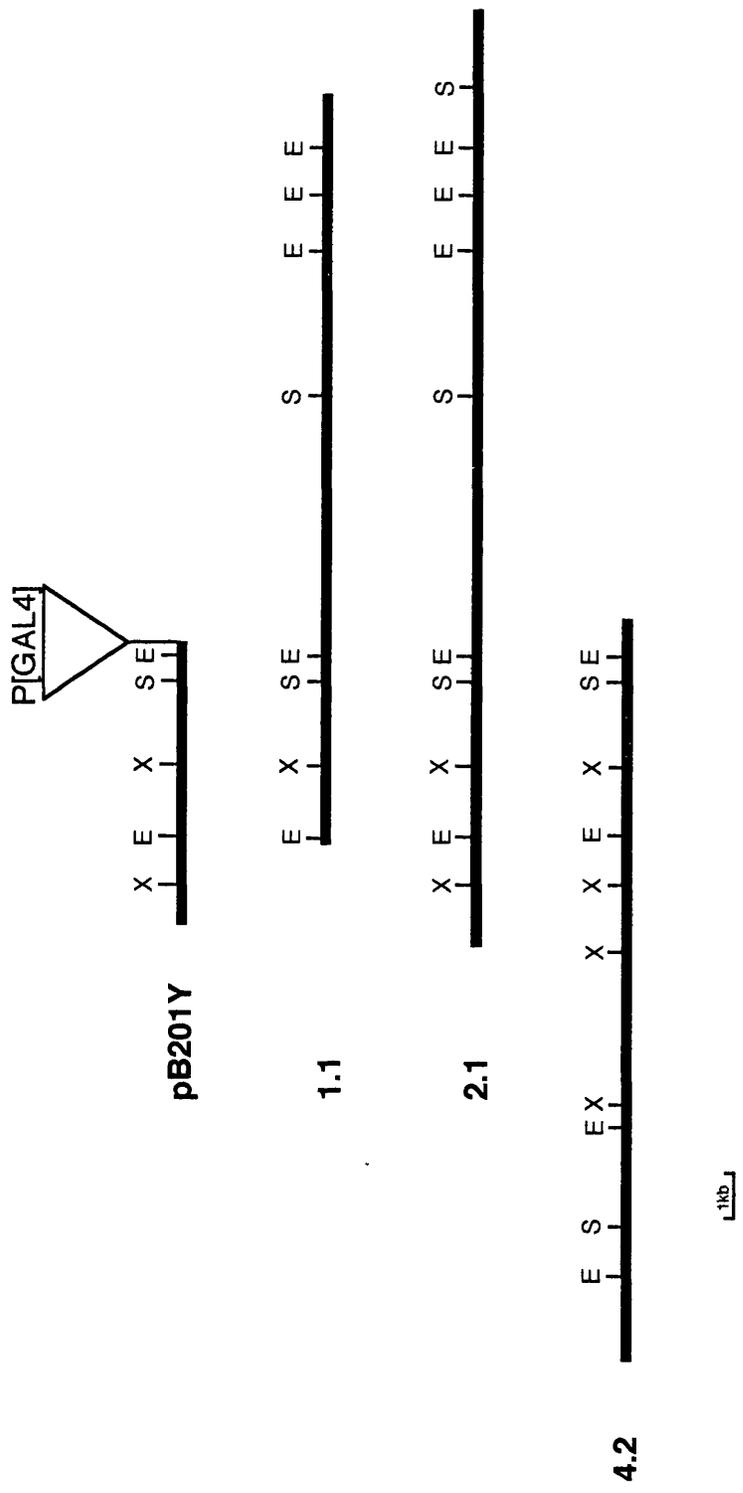
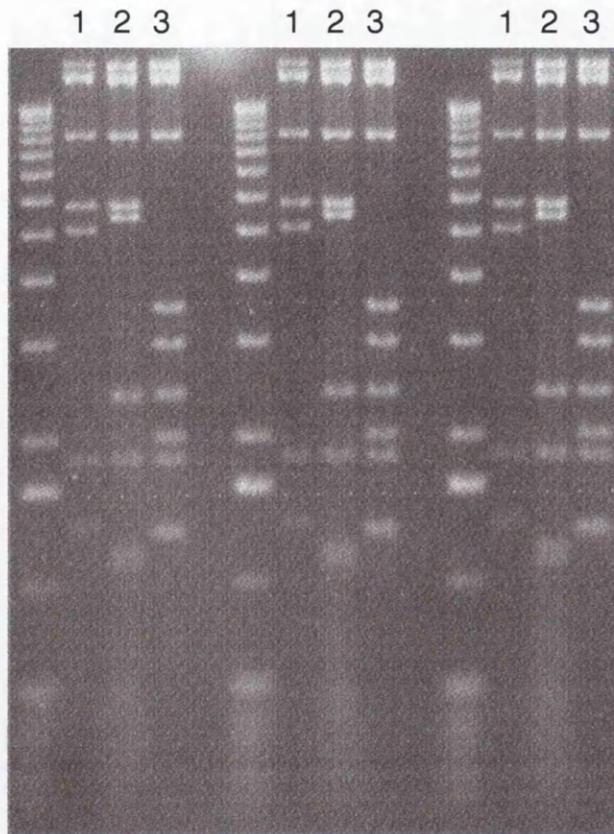


Figure 5.4. Restriction maps of genomic clones 1.1, 2.1, and 4.2, and the rescued plasmid pB201Y. E = *EcoRI*, X = *XhoI*, S = *SstI*.



**Figure 5.5.** *XhoI* + *SstI* double restriction digests of genomic clones 1.1, 2.1, and 4.2. Lane 1 = clone 1.1, lane 2 = clone 2.1, lane 3 = clone 4.2. Each set of three digests was Southern blotted onto separate filters.

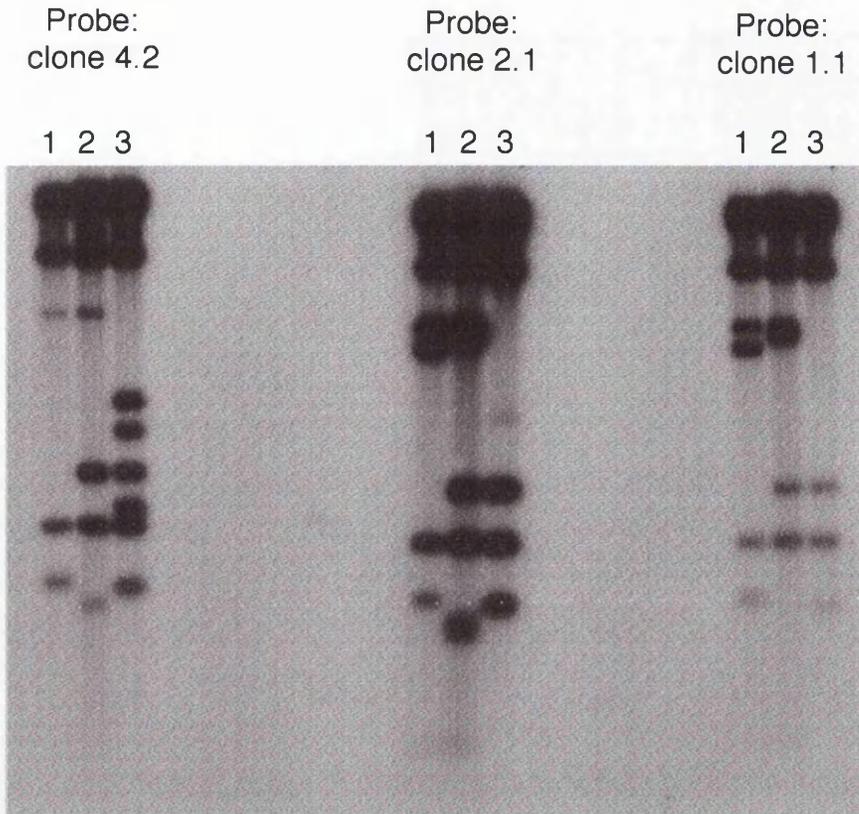
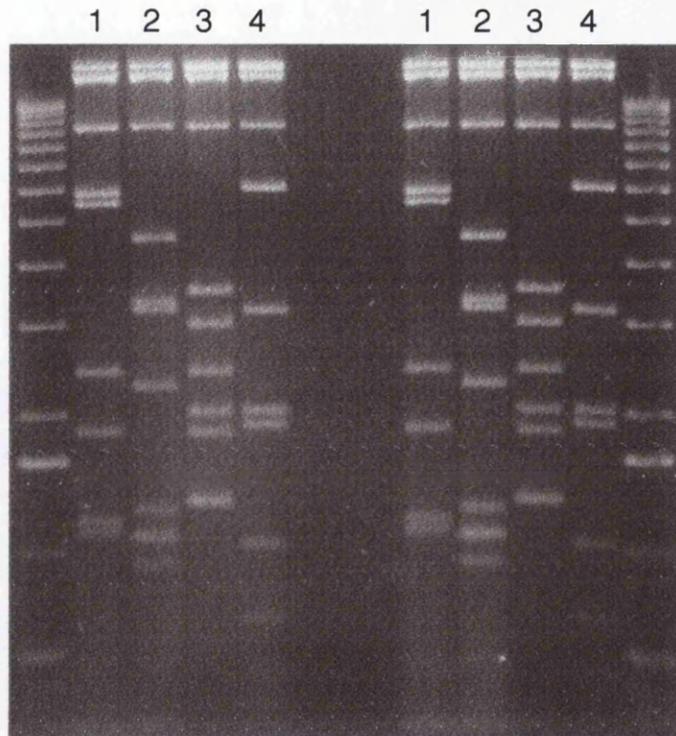
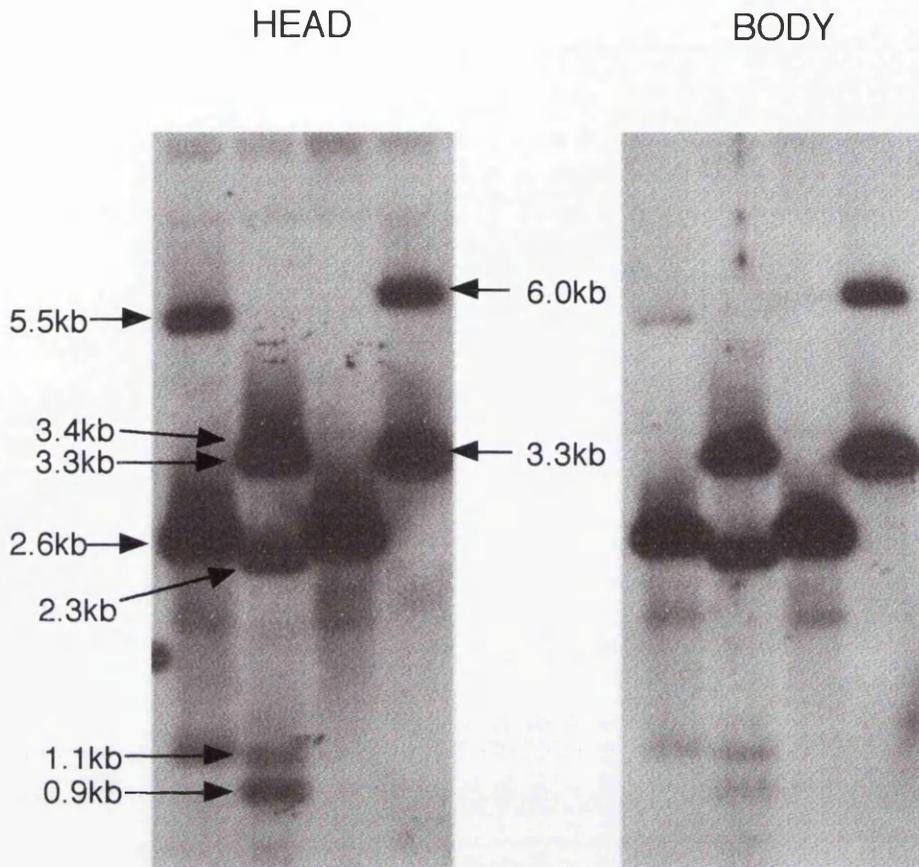


Figure 5.6. Gels from figure 5.5 Southern blotted and probed with whole phage as indicated. Lane 1 = clone 1.1, lane 2 = clone 2.1, lane 3 = clone 4.2.



**Figure 5.7.** Restriction digests of genomic clones 2.1 and 4.2. Lanes 1 and 2 = clone 2.1, lanes 3 and 4 = clone 4.2. Lanes 1 and 3 = *XhoI* + *SstI*, lanes 2 and 4 = *SstI* + *EcoRI*. Both sets of digests blotted separately and used in reverse Northern analysis (figure 5.8).



**Figure 5.8.** Reverse Northern analysis of genomic clones 2.1 and 4.2. Gels in figure 5.7 blotted and probed separately with head cDNA probe or body cDNA probe. Sizes and positions of fragments discussed in the text are indicated.

discussed in the previous chapter, i.e. the 2.4kb *XhoI/SstI* band, the 3.3kb *SstI/EcoRI* band, the 2.3kb *SstI/EcoRI* band of clone 2.1, and the 6.0kb *SstI/EcoRI* band of clone 4.2. However, the 5.5kb *XhoI/SstI* band of clone 2.1 appears to contain a newly identified transcriptional unit. Furthermore, it clearly has elevated expression in the head, as the hybridisation is far stronger to the head probe than the body probe. Obviously, *SstI/EcoRI* fragments from clone 2.1 should also contain this transcriptional unit. Hybridisation occurs to the 3.4kb *SstI/EcoRI* fragment (but this is largely obscured by the strong hybridisation to the 3.3kb fragment), and to the 1.1 and 0.9kb fragments.

### 5.2.3 Isolation and Sequence analysis of cDNA clones

To identify this new transcriptional unit it was decided to screen a *Drosophila* head cDNA library constructed in lambda phage NM1149. Initially 100,000 recombinant phage were plated out and duplicate filter lifts probed with 5.5kb *SstI* fragment of genomic clone 2.1. Six positives were identified and purified down to single plaques. Phage DNA was prepared and restriction digested with *EcoRI* and *HindIII*. Restriction fragments were separated by gel electrophoresis and results are shown in figure 5.9. To confirm the identified phage as real positives, the gel was Southern blotted and probed with the 5.5kb *SstI* fragment of genomic clone 2.1. The result is also shown in figure 5.9, and as can be seen only clones A421, A231, and A911 hybridised to the probe. As clones A522, A341, and A422 did not hybridise to the probe, further analysis of these clones was not pursued. Furthermore, it is clear that clones A421, A231 and A911 represent the same cDNA, and therefore further analysis was restricted to A421.

Prior to sequencing and further analysis of the A421 cDNA, it was decided to subclone restriction fragments from it into the plasmid vector pBluescript SK<sup>-</sup>. Initially, phage DNA from clone A421 was digested with *EcoRI*, *HindIII*, and *EcoRI* and *HindIII*, and the *EcoRI/HindIII* fragments subcloned into pBluescript, generating

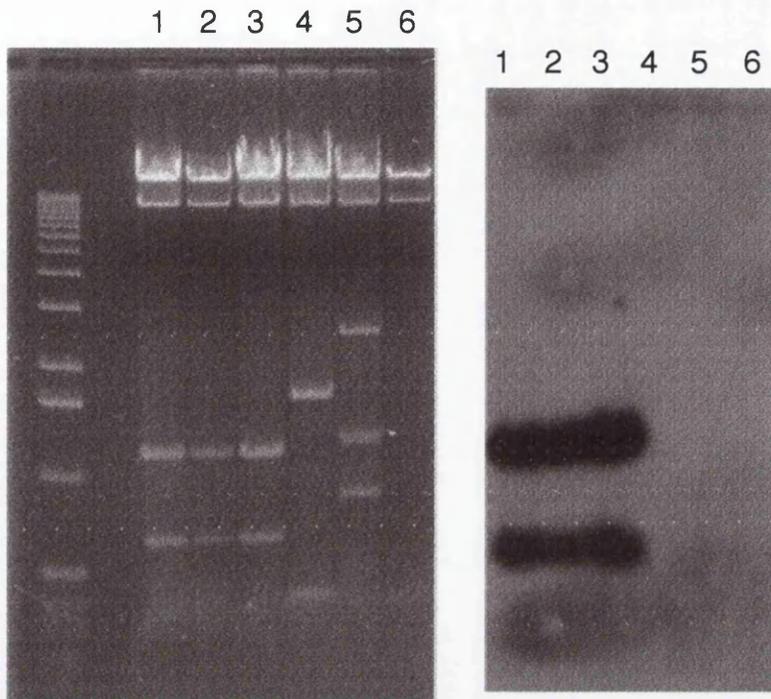
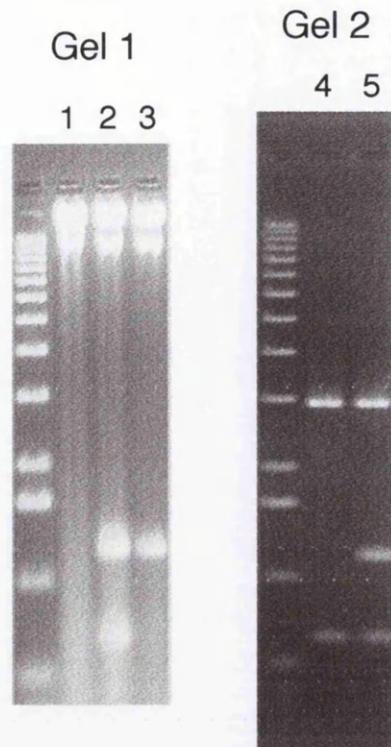


Figure 5.9. Isolated cDNA clones digested with *EcoRI* and *HindIII*. Lane 1 = A421, lane 2 = A231, lane 3 = A911, lane 4 = A522, lane 5 = A341, lane 6 = A422. Gel Southern blotted and probed with the 5.5kb *SstI* fragment of genomic clone 2.1. Result is shown on the right. Only clones A421, A231, and A911 hybridise to the probe.

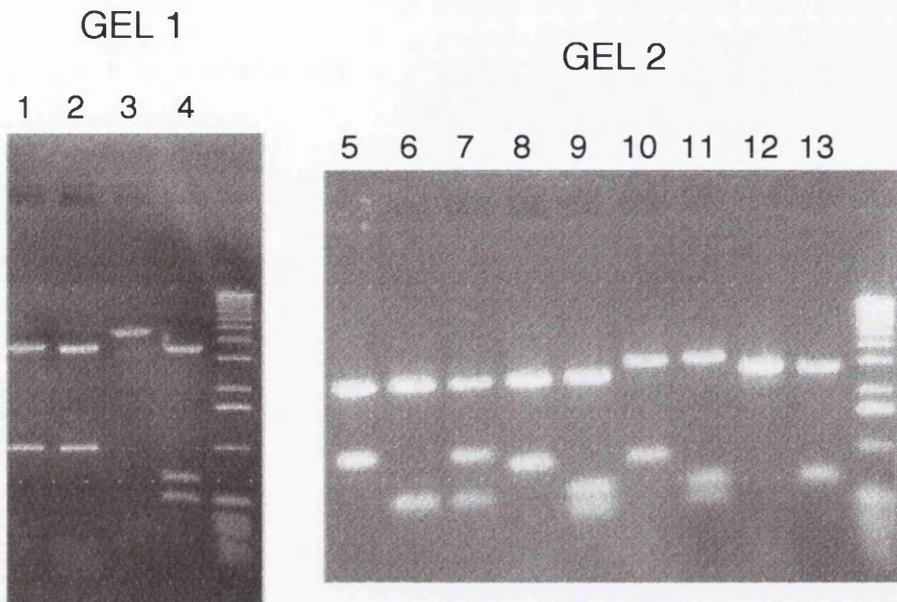
new recombinant plasmids pAMA1 and pAMA2 (figure 5.10). Note that plasmid pAMA1 contains the entire A421 cDNA, and further subclones were generated from this plasmid rather than the phage cDNA. Plasmid pAMA1 was subject to further restriction digests, i.e. *PstI*, *PstI* and *EcoRI*, *EcoRV*, *EcoRV* and *EcoRI*, and some of the fragments generated further subcloned into pBluescript (figure 5.11). Figure 5.12 summarises the plasmid subclones generated from cDNA A421.

Each of the plasmids pAMA2, pAMA4, and pAMA5 was sequenced on both strands, initially using primers T3 and T7, and generating further sequence by using synthesised oligonucleotides as primers, where necessary. Plasmid pAMA3 was also sequenced on both strands, but only for the region not overlapping with pAMA4 and pAMA5. From the sequences generated a single contig of 1808 bases was constructed. This contig has a single large ORF of 474 amino acids (figure 5.13).

It would appear that A421 was not a full-length cDNA as the ORF does not begin with a translation start codon. Further 5' sequence was obtained from the genomic clone 2.1. The 5.5kb *SstI* fragment of 2.1 was subcloned into pBluescript, and a oligonucleotide primer designed, from the known sequence of A421, to obtain further sequence. The genomic sequence obtained was combined with the cDNA sequence of A421 to give the sequence of 1986 nucleotides shown in figure 5.14. This combined sequence has a largest ORF of 510 amino acids, starting with a translational start codon, and having an in frame stop codon 12bp upstream. However, as the first 188 nucleotides of this sequence was obtained from genomic DNA, there must be some question as to whether this translation start codon is actually used. The consensus sequence for *Drosophila* translation initiation sites has been studied by Cavener, (1987). From a study of 100 *Drosophila* genes he derived the following consensus sequence: C/A(57/32) A(82) A(56) A/C(38/36) ATG, figures in brackets are observed percentage frequencies for each base. The corresponding genomic sequence is A A A C ATG, which is a good match to the consensus, suggesting the identified ATG codon may be a good candidate for translation initiation. Of course, the possibility of an intron existing within the genomic sequence cannot be ruled out. However, there is no



**Figure 5.10.** Subcloning of cDNA clone A421 restriction fragments. Gel 1: restriction digests of cDNA clone A421. Lane 1 = *HindIII*, lane 2 = *EcoRI* + *HindIII*, lane 3 = *EcoRI*. Gel 2: subcloned fragments. Lane 4 = pAMA2 *EcoRI* + *HindIII* digested, lane 5 = pAMA1 *EcoRI* + *HindIII* digested.



**Figure 5.11.** Subcloning of pAMA1 restriction fragments. Gel 1: restriction digests of pAMA1. Lane 1 = *PstI*, lane 2 = *PstI* + *EcoRI*, lane 3 = *EcoRV*, lane 4 = *EcoRV* + *EcoRI*. Gel 2: plasmids containing pAMA1 restriction fragments. Plasmids and restriction digests are as follows: Lane 5 = pAMA3, *EcoRI*, lane 6 = pAMA2, *HindIII*, lane 7 = pAMA1, *EcoRI* + *HindIII*, lane 8 = pAMA5, *EcoRI* + *PstI*, lane 9 = pAMA5, *EcoRI* + *EcoRV*, lane 10 = pAMA1, *PstI* + *EcoRI*, lane 11 = pAMA1, *EcoRI* + *EcoRV*, lane 12 = pAMA4, *PstI* + *EcoRI*, lane 13 = pAMA5, *EcoRI* + *EcoRV*.

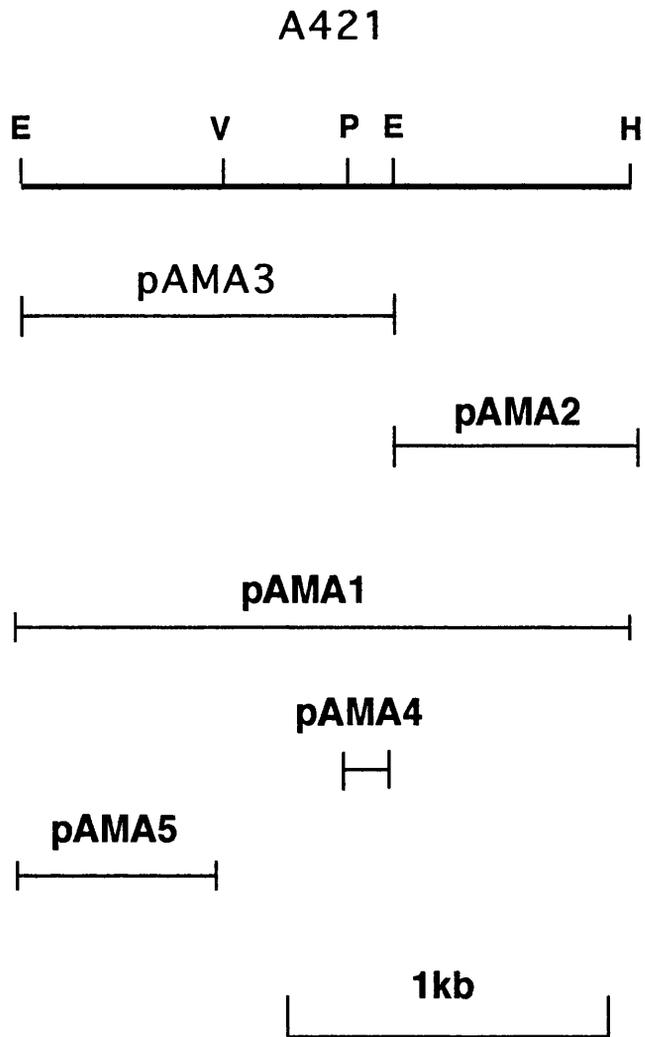


Figure 5.12. A421 restriction fragment subclones. Restriction map of A421 cDNA is shown at the top. Horizontal bars represent regions of A421 sub-cloned into indicated plasmids. E = *EcoRI*, V = *EcoRV*, P = *PstI*, H = *HindIII*

A AGC CGG CTG CTG TCA CCA AGT GCG CCT GGT GGA  
 S R L L S P S A P G G  
 50  
 CGC AAG AGC TTC ACC TCG CTC AAT CTC ACC CTG CGC CAG CCG ACG GGC  
 R K S F T S L N L T L R Q P T G  
 100  
 TCG GCA CAA TCG GCC ATT GAC ATC ACC GCC GGT CCG GCG CCG AGT GGG  
 S A Q S A I D I T A G P A P S G  
 150  
 CAG GGC AGT GGG ATC ACC TAC TCC AGC GTT AGT TTT GAC GCA CGT CGT  
 Q G S G I T Y S S V S F D A R R  
 200  
 GGG ACG CAG AAG AAT TTC CAG CTG ACT GTG ACG GAC GAG GGC AGT GTC  
 G T Q K N F Q L T V T D E G S V  
 250  
 TTC AGT GCG GGC AGC GTT CGC CCA AGG ACC TTG TAC GCA ACG CCC TGT  
 F S A G S V R P R T L Y A T P C  
 300  
 GAG CCG GTG ACA AGT GTT CCG GCA AAC CAG CAG CCT CCT CCG CCC GCA  
 E P V T S V P A N Q Q P P P P A  
 350  
 GCA GTG GTG GCA GAT GGT CTG GGG GAC GCT GAT ATG ACG TCG GAC GTT  
 A V V A D G L G D A D M T S D V  
 400  
 TTC CCA TAT CCC ACG CAG GAG CAG AAT CAC ATT GTG GCA TCC AAC TAT  
 F P Y P T Q E Q N H I V A S N Y  
 450  
 AGT AAC AAC CAT GCC GTC AGC GAC AAT AAC AAC ATT GGG ATC AGC AAT  
 S N N H A V S D N N N I G I S N  
 500  
 AGC ACG GAC AGT AGT CCC ACT ATG CCC CTT TAC GAA GGT GTC ATG GAG  
 S T D S S P T M P L Y E G V M E  
 550  
 GAA TGC GAT CGC GAA CGT CGT GCC GCC ACT ATC GAG CGT CAG AAG CAG  
 E C D R E R R A A T I E R Q K Q  
 600  
 CGT CGC GAC AAG CTG GCC AAT GCA CTG AGG GAC AAC AAG AAG CGT CTG  
 R R D K L A N A L R D N K K R L  
 650  
 CTC GTC CTG GAG CAG GAG ATC AAC ATC CTG ACT GAG CCG GTG CCG GTG  
 L V L E Q E I N I L T E P V P V  
 700  
 GGG GAA TCT GAA AGA CTG GAT AGA GAT ATC AAG CAA CTG ACT GAG GAC  
 G E S E R L D R D I K Q L T E D  
 750  
 AAT CGG CTG CTT GAC TGC TTA AAT GAA CCG TGT CAG GCG AAT GGC CCT  
 C N R L L D C L N E P Q A N G P  
 800  
 GGT CCG GCG GCC AAT CCC ACG AAC CGC CAG CAC CTG TTG CCA GGA AGC  
 G P A A N P T N R Q H L L P G S  
 850  
 AAT GCT TCA CAG CAG CAG CCG CAG CAA CAG CCA CAG CCA TTC CCG CGT  
 N A S Q Q Q P Q Q Q P Q P F P R  
 900  
 CAG CGT CAG AGT GCT CGC GTC CAG GCG CCG CCC AGT TCA CTA CGT CTG  
 Q R Q S A R V Q A P P S S L R L  
 CAC TCG GTG CCG GCA GCA CCC AAC TCC CAG CCG GTG CAG GAC TTT GGG  
 H S V P A A P N S Q P V Q D F G

950  
 CAG CAC CAC AGC AGT GCT CCC ACC TCG GCC TGT TTA ACG CCC CAG ATG  
 Q H H S S A P T S A C L T P Q M  
 1000  
 CAG AGT CAA CAG CAA CTG TTA CAG CTG CAG CAG GAG CCG CCA CCG ACG  
 Q S Q Q Q L L Q L Q Q E P P P T  
 1050  
 TAC GCC CAG TAC TAT CAG TTC CAG CAA TAT CTG CAG CAG CAG CGA CAG  
 Y A Q Y Y Q F Q Q Y L Q Q Q R Q  
 1100  
 CAG CAG TTA CAA CAG CAG ATG CAG CAG CAA CAG CAG CTA CAA CAG CAG  
 Q Q L Q Q Q M Q Q Q Q Q L Q Q Q  
 1150  
 ATG CAG CAG CAA CAG CAG CCG CAG CCA CAA CAG GAG GAA GAA TTC CTG  
 M Q Q Q Q Q P Q P Q Q E E E F L  
 1200  
 TCC GAC TCC GAT GTG GAT GAG GAG GAG GAG ACT CTG GAC TCG TGG GCC  
 S D S D V D E E E E T L D S W A  
 1250  
 TGC AAC ATG TGT ACA TTC CGC AAC CAT CCA CAG CTA AAT ATT TGT GAG  
 C N M C T F R N H P Q L N I C E  
 1300  
 GCC TGT GAG AAC GTT AGG ATC CAG CCG GGT ATG ATA CGC ATT GTT CCC  
 A C E N V R I Q P G M I R I V P  
 1350  
 AGT GGA GGC GGG GCC GCT GCT GCT GCT GCC ACG CCA CCC GGC AGC ATT  
 S G G G A A A A A A T P P G S I  
 1400  
 GAG CAG CAA CAG CAG CCG TCC CAG CAG CCG TAC GCT CTG CAT ACA TAA  
 E Q Q Q Q P S Q Q P Y A L H T \*  
 1450  
 CCCAAAAGCTGCACCAAACGATAACAAAGTTCTCTTATCGTGACTGTTGTTCTTGCATGCTCG  
 1500 1550  
 ACTTTGAAACTGTGATAGCATTTAAACTAGTAGGCGGCAGGTTTATTATCACCTGCATGGCTCC  
 1600  
 ACCTGAAAGACATGTGCGAAACGAAAACAACCGGAACGAAACGAATTGCTTAAGCGCGGCCTT  
 1650  
 TTAAATTATTTAACATATTCGATTTATGTATGTACATTAATGCAATCGCCTCAGCCTCAGG  
 1700  
 CCGTTTAATTTAACAATATCCTCTAACCCTTAAACTAGCATTAAAGCCAACAAGAAACAGAGTTT  
 1750  
 ATTATTTTTAGCCCGCTCAAAGTAAATAAAGAACGTTATTTAAATATTTAAAAA

Figure 5.13. Complete DNA sequence of cDNA clone A421, and translation of largest ORF. Polyadenylation signal is underlined.

good match to the consensus for *Drosophila* splice donor sites of MAGGTRAGTW (underlined residues absolutely conserved) worked out by Mount *et al.* (1992). But there is a reasonable match to the splice acceptor site consensus, of TTTTTYYYTTNCAGRT (bases before the N are not highly conserved), between bases 61-76 of the sequence in figure 5.14. It is entirely possible that this acceptor site is used in other splice variants of the gene.

The peptide sequence from figure 5.14 was used to search protein sequence databases using the BlastP algorithm. No significant homology to sequences in the databases could be found, suggesting this is a new gene of unknown function. A search of the Prosite database of protein sequence motifs also found no matches.

#### **5.2.4 Genomic Southern**

To determine if the A421 gene exists as a single copy within the genome of *Drosophila* a genomic Southern was performed. Genomic DNA from wild-type Oregon-R flies was restriction digested with *EcoRI*, *SstI*, and *XhoI*, and separated by gel electrophoresis, together with genomic DNA from 201Y flies digested with *SstI* and *XhoI*. The gel was Southern blotted onto nylon filters and probed with the whole pAMA1 plasmid (figure 5.15). Within the wild-type DNA a single *SstI* band of 5.5kb is evident, suggesting that the A421 gene is single-copy. Two *EcoRI* bands of 8.5kb and 1.0kb are present, consistent with the A421 gene straddling the *EcoRI* site 8.0kb upstream from the P[GAL4] insertion site. No distinct *XhoI* band is visible, probably due to the extreme size of this band, as no *XhoI* site is present within the region upstream of the P[GAL4] element covered by the genomic clones. The 201Y genomic DNA gives a 5.5kb *SstI* band consistent with the result from wild-type genomic DNA. The other bands present in the 201Y lanes, i.e. the 3.6kb *SstI* band and the 5.2kb *XhoI* band, are generated from sequences within the P[GAL4] element which hybridise to the pBluescript backbone of the whole plasmid probe.

AAGTGCACAGGTGTGGGTGGATCTGCTCAGTGTGCAAATAACAAATTGTGAAAATAACTCGCA

100  
AAC ATG CAG ATA CAC CTA GTG CAC CAA ATA AAG CAA AGA TCA ACA CAT  
M Q I H L V H Q I K Q R S T H

150  
TTC CAA ACC GGG CAA CAA GTT GTG AAG CGT TTT TTA AGG GGA GGG GCC  
F Q T G Q Q V V K R F L R G G A

200  
ATC CCA AGT TTC CCA AGC CGG CTG CTG TCA CCA AGT GCG CCT GGT GGA  
I P S F P S R L L S P S A P G G

250  
CGC AAG AGC TTC ACC TCG CTC AAT CTC ACC CTG CGC CAG CCG ACG GGC  
R K S F T S L N L T L R Q P T G

300  
TCG GCA CAA TCG GCC ATT GAC ATC ACC GCC GGT CCG GCG CCG AGT GGG  
S A Q S A I D I T A G P A P S G

350  
CAG GGC AGT GGG ATC ACC TAC TCC AGC GTT AGT TTT GAC GCA CGT CGT  
Q G S G I T Y S S V S F D A R R

400  
GGG ACG CAG AAG AAT TTC CAG CTG ACT GTG ACG GAC GAG GGC AGT GTC  
G T Q K N F Q L T V T D E G S V

450  
TTC AGT GCG GGC AGC GTT CGC CCA AGG ACC TTG TAC GCA ACG CCC TGT  
F S A G S V R P R T L Y A T P C

500  
GAG CCG GTG ACA AGT GTT CCG GCA AAC CAG CAG CCT CCT CCG CCC GCA  
E P V T S V P A N Q Q P P P P A

GCA GTG GTG GCA GAT GGT CTG GGG GAC GCT GAT ATG ACG TCG GAC GTT  
A V V A D G L G D A D M T S D V

550  
TTC CCA TAT CCC ACG CAG GAG CAG AAT CAC ATT GTG GCA TCC AAC TAT  
F P Y P T Q E Q N H I V A S N Y

600  
AGT AAC AAC CAT GCC GTC AGC GAC AAT AAC AAC ATT GGG ATC AGC AAT  
S N N H A V S D N N N I G I S N

650  
AGC ACG GAC AGT AGT CCC ACT ATG CCC CTT TAC GAA GGT GTC ATG GAG  
S T D S S P T M P L Y E G V M E

700  
GAA TGC GAT CGC GAA CGT CGT GCC GCC ACT ATC GAG CGT CAG AAG CAG  
E C D R E R R A A T I E R Q K Q

750  
CGT CGC GAC AAG CTG GCC AAT GCA CTG AGG GAC AAC AAG AAG CGT CTG  
R R D K L A N A L R D N K K R L

800  
CTC GTC CTG GAG CAG GAG ATC AAC ATC CTG ACT GAG CCG GTG CCG GTG  
L V L E Q E I N I L T E P V P V

850  
GGG GAA TCT GAA AGA CTG GAT AGA GAT ATC AAG CAA CTG ACT GAG GAC  
G E S E R L D R D I K Q L T E D

900  
AAT CGG CTG CTT GAC TGC TTA AAT GAA CCG TGT CAG GCG AAT GGC CCT  
C N R L L D C L N E P Q A N G P

950  
GGT CCG GCG GCC AAT CCC ACG AAC CGC CAG CAC CTG TTG CCA GGA AGC  
G P A A N P T N R Q H L L P G S

1000  
 AAT GCT TCA CAG CAG CAG CCG CAG CAA CAG CCA CAG CCA TTC CCG CGT  
 N A S Q Q Q P Q Q Q P Q P F P R  
 1050  
 CAG CGT CAG AGT GCT CGC GTC CAG GCG CCG CCC AGT TCA CTA CGT CTG  
 Q R Q S A R V Q A P P S S L R L  
 1100  
 CAC TCG GTG CCG GCA GCA CCC AAC TCC CAG CCG GTG CAG GAC TTT GGG  
 H S V P A A P N S Q P V Q D F G  
 1150  
 CAG CAC CAC AGC AGT GCT CCC ACC TCG GCC TGT TTA ACG CCC CAG ATG  
 Q H H S S A P T S A C L T P Q M  
 1200  
 CAG AGT CAA CAG CAA CTG TTA CAG CTG CAG CAG GAG CCG CCA CCG ACG  
 Q S Q Q Q L L Q L Q Q E P P P T  
 1250  
 TAC GCC CAG TAC TAT CAG TTC CAG CAA TAT CTG CAG CAG CAG CGA CAG  
 Y A Q Y Y Q F Q Q Y L Q Q Q R Q  
 1300  
 CAG CAG TTA CAA CAG CAG ATG CAG CAG CAA CAG CAG CTA CAA CAG CAG  
 Q Q L Q Q Q M Q Q Q Q Q L Q Q Q  
 1350  
 ATG CAG CAG CAA CAG CAG CCG CAG CCA CAA CAG GAG GAA GAA TTC CTG  
 M Q Q Q Q Q P Q P Q Q E E E F L  
 1400  
 TCC GAC TCC GAT GTG GAT GAG GAG GAG GAG ACT CTG GAC TCG TGG GCC  
 S D S D V D E E E E T L D S W A  
 1450  
 TGC AAC ATG TGT ACA TTC CGC AAC CAT CCA CAG CTA AAT ATT TGT GAG  
 C N M C T F R N H P Q L N I C E  
 1500  
 GCC TGT GAG AAC GTT AGG ATC CAG CCG GGT ATG ATA CGC ATT GTT CCC  
 A C E N V R I Q P G M I R I V P  
 1550  
 AGT GGA GGC GGG GCC GCT GCT GCT GCT GCC ACG CCA CCC GGC AGC ATT  
 S G G G A A A A A A T P P G S I  
 1600  
 GAG CAG CAA CAG CAG CCG TCC CAG CAG CCG TAC GCT CTG CAT ACA TAA  
 E Q Q Q Q P S Q Q P Y A L H T \*  
 1650  
 CCCAAAAGCTGCACCAAACGATAACAAAGTTCTCTTATCGTGACTGTTGTTCTTGCATGCTCG  
 1700  
 ACTTTGAAACTGTGATAGCATTAAACTAGTAGGCGGCAGGTTTATTATCACCTGCATGGCTCC  
 1750 1800  
 ACCTGAAAGACATGTGCGAAACGAAAACAACGCGAACGGAAACGAATTGCTTAAGCGCGCCTT  
 1850  
 TTTAAATTATTTTAAACATATTTCGATTTATGTATGTACATTAAATGCAATCGCCTCAGCCTCAGG  
 1900  
 CCGTTTAATTTAACAATATCCTCTAACCGTTAAACTAGCATTAAAGCCAACAAGAAACAGAGTTT  
 1950  
 ATTATTTTTAGCCCGCTCAAAAGTAAATAAAGAACGTTATTTAAATATTTAAAA

Figure 5.14. A421 cDNA sequence plus additional 5' genomic sequence. End of the genomic sequence is indicated by \* above the base. Translation of largest ORF is given beneath the DNA sequence. In frame upstream stop codon is double underlined. Sequence conforming to *Drosophila* splice acceptor site consensus is underlined. Polyadenylation signal is in bold.

### 5.2.5 Expression of the A421 gene

Northern analysis reveals four transcripts produced by the A421 gene. Figure 5.16 shows the result of probing equal amounts of head and body polyA<sup>+</sup> RNA, from wild-type Oregon-R flies and fly line 201Y, with the cDNA from pAMA1. Four transcripts of 1.6, 2.0, 2.6, and 3.3kb are evident, with the 2.0kb transcript being clearly head elevated. The head elevated transcript is by far the most abundant transcript, which is consistent with the reverse Northern analysis of the genomic clones which identified a clearly head elevated transcriptional unit. No significant difference between transcripts in wild-type flies and 201Y flies is evident, indicating that the 201Y P[GAL4] insertion does not disrupt the expression of this gene. This may have been expected as the P[GAL4] element is approximately 6kb from the region containing the A421 gene, and therefore is unlikely to have disrupted coding regions or control elements of this gene.

Presumably the A421 cDNA corresponds to the 2.0kb transcript because of the similarity in sizes, and because the A421 cDNA was isolated from a head cDNA library in which the 2.0kb transcript must have predominated. The longer transcripts are possibly derived from other, as yet unidentified, exons spliced on to the potential splice acceptor site discussed earlier.

To look at the expression of the A421 gene further, *in situ* hybridisations to fly head sections were performed, using DIG labelled probes. Using a T7 primer, single-stranded DIG labelled anti-sense probes were generated by thermal cycling from plasmid pAMA1, as outlined in Chapter 2. The result of hybridising this probe to fly head sections is shown in figure 5.17. The cell bodies of most of the intrinsic neurons of the *Drosophila* brain exist in a thin layer over the surface of the brain. As can be seen in figure 5.17, the probe hybridises to this layer in a characteristic pattern, suggesting that the A421 gene is expressed in all, or most, of the neurons of the *Drosophila* brain. In particular, hybridisation is not restricted to the Kenyon cell bodies, which have a dorsal and posterior location to the mushroom body calyx, and are located in small regions at the top of the fly brain in the section shown. A control

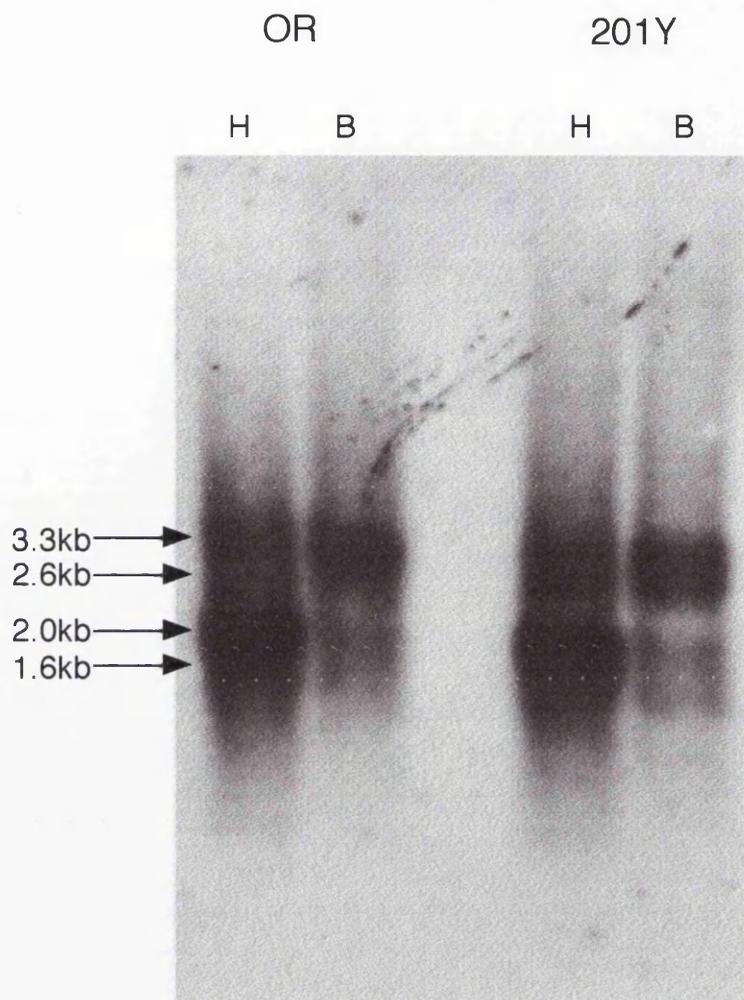


Figure 5.16. Northern blot of poly A<sup>+</sup> RNA from Oregon-R (OR) or 201Y fly heads (H) or bodies (B), probed with pAMA1 cDNA. Sizes of bands discussed in the text are indicated.

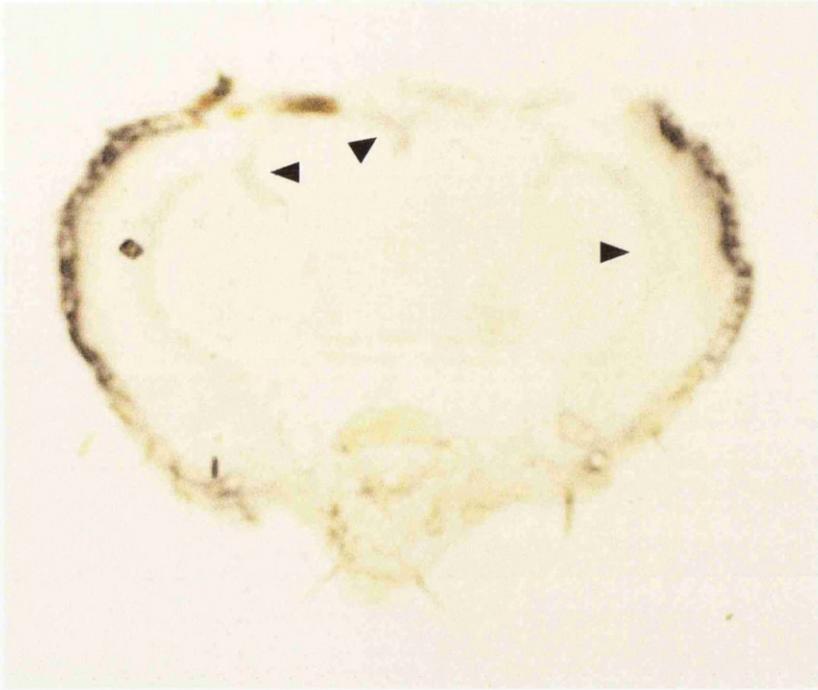


Figure 5.17 Frontal head section of wild-type adult female fly probed with DIG labelled cDNA A421, and stained for alkaline phosphatase activity. Staining can be seen in regions where CNS cell bodies are known to reside (indicated by arrows) but is not restricted to the mushroom bodies.

sense probe, generated from pAMA1 with a T3 primer, showed no hybridisation to head sections (data not shown). Therefore, it must be concluded that possible enhancers that control A421 gene expression are probably not responsible for 201Y enhancer-trap expression.

### 5.3 Discussion

As mentioned in section 5.2.3 the protein sequence of gene A421 has no significant homology to any sequences in the databases, nor does it contain any recognised protein motifs. However, a closer analysis of the protein sequence reveals a number of features that are suggestive of a potential function. Analysis of these features is now presented. Figure 5.18 gives the A421 protein sequence with the main features discussed here highlighted.

#### 5.3.1 A421 has a Glutamine-Rich Domain.

Perhaps the most obvious feature of the A421 protein is the glutamine-rich region between amino acids 366 and 426 (figure 5.18). 37 of the 61 amino acids of this region are glutamines, with a particularly glutamine-rich central region, between amino acids 395 and 421, where 22 of the 27 amino acids are glutamines. It has been known for some time that glutamine-rich regions or polyglutamine stretches can act as transcriptional activation domains in some transcription factors (reviewed in Mitchell and Tjian, 1989). One of the first homopolymeric glutamine repeats to be characterised was found in the *Drosophila* transcription factor *Antennapedia* (Schneuwly *et al.*, 1986). The glutamine-rich region of *Antennapedia* protein can activate transcription, in *Drosophila* SL2 tissue culture cells, when fused to the DNA binding domain of the human transcription factor Sp1 (Courey *et al.*, 1989). The human transcription factor Sp1 has itself a glutamine-rich domain which has been, by deletion analysis, shown to be essential for transcriptional activation in a *Drosophila* tissue culture system (Courey

1 MQIHLVHQIK QRSTHFQTGQ QVVKRFLRGG AIPFSPSRL L SPSAPGGRKS FTSLNLTLRQ  
61 PTGSAQSAID ITAGPAPSGQ GSGITYSSVS FDARRGTQKN FQLTIVTDEGS VFSAGSVRPR  
121 TLYATPCEPV TSVPANQQPP PPAAVVADGL GDADMTSDVF PYPTQEQNHI VASNYSNNHA  
181 VSDNNNIGIS NSTDSSPTMP LYEGVMEECD RERRAATIER QKQRRDKLAN ALRDNKKRLL  
241 VLEQEINILT EPVPVGESE LDRDIKOLTE DCNRLLDCLN EPQANGPGPA ANPTNRQHLL  
301 PGSNASQQQP QQQPQFFPRQ RQSARVQAPP SSLRLHSVPA APNSQPVQDF GQHHSSAPTS  
361 ACLTPQMOSQ QQLLQLOQEP PPTYAQYYQF QQYLQQRQQ QLQQQMQQQQ QLQQQMQQQQ  
421 QPOQOQEEEF LSDSVDDEEE ETLDSWACNM **CTFRNHPQLN** **ICEACENVRI** QPGMIRIVPS  
481 GGGAAAAAAT PPGSIEQQQQ PSQQPFSLHT

**Figure 5.18.** A421 amino acid sequence. Gln-rich regions are double underlined. Region rich in acidic residues is single underlined. Region containing putative novel Zn-finger-like motif is in bold. Possible nuclear localisation signal marked \*\*\*\*\*. Leucines involved in potential leucine-zipper highlighted in bold.

and Tjian, 1988). Simple homopolymeric stretches of at least six glutamines can activate transcription in HeLa cells when fused to the DNA binding domain of the transcription factor GAL4. Similarly, the glutamine-rich domains of transcription factors OCT-2A, Spl, and TATA binding protein, could all activate transcription when fused to the DNA binding domain of GAL4 (Gerber *et al.*, 1994). The glutamine-rich regions of the *Drosophila* transcription factor *twist* can also activate transcription in yeast when fused to the GAL4 DNA binding domain (Chung *et al.*, 1996).

A search of the Genbank database supports the evidence that glutamine-rich regions play an important role in transcription factors. Figure 5.19 is a list of all *Drosophila* proteins (up to June 1995) in the Genbank database having domains described as glutamine-rich or polyglutamine by the submitting authors. All the proteins classified as transcription factors have recognised DNA binding motifs, such as homeobox domains, zinc-fingers, leucine-zippers etc. The proteins classified as developmental regulators are either expressed only at specific developmental stages, or are from loci linked with developmental mutations. Of the 54 proteins listed 38, or 70%, are transcription factors. Gerber *et al.*, (1994) screened the SwissProt protein database for proteins from any species containing polymeric stretches of at least 20 glutamines, and found that 82% of the 40 top-scoring proteins were transcription factors. These data clearly demonstrate that glutamine-rich or poly-glutamine stretches occur predominantly in transcription factors.

Within the glutamine-rich region of A421 there are a number of amino acids with large hydrophobic sidechains, particularly leucine and methionine. The glutamine-rich region of the human transcription factor SpI also contains a number of large hydrophobic residues, and substitution of these greatly reduced its interaction with the dTAF<sub>II</sub>110 component of the *Drosophila* TFIID TATA-box-binding complex, and its ability to activate transcription (Gill *et al.*, 1994). Interestingly, deletion of large hydrophobic residues in the Q2 activation domain of the transcription factor CREB also greatly reduced its interaction with dTAF<sub>II</sub>110 (Ferreri *et al.*, 1994). It would appear that alternating hydrophobic and polar, in particular glutamine, residues may be

Transcription factors:

abdominal-a  
abdominal-b  
antennapedia  
bicoid  
brahma  
cap 'n' collar  
cf2 (I and II)  
cf2 III  
cut  
deformed  
ecdysone induced protein  
empty spiracles  
engrailed  
extra-machrochaetae  
giant protein  
ftz-f1  
grainy head  
hairy  
hunchback  
knirps  
labial  
odd-paired  
odd-skipped  
proboscipedia  
prospero  
proximal chromatin protein  
rough  
sex combs reduced  
single-minded  
slow border cell protein  
suppressor of hairless  
suppressor of zeste  
tfII-d  
twist  
wings-down  
yan  
zeste  
zfh-1

Developmental  
regulators:

big brain  
centrosomal mitotic factor  
elav  
eyes absent  
mastermind  
maternal pumilio protein  
mitosis initiation protein  
nanos  
notch  
small optic lobes  
staufen

Miscellaneous:

diacylglycerol kinase  
ring canal protein  
rutabaga  
shakB  
suppressor of sable  
tyrosine phosphatase

Figure 5.19. List of *Drosophila* proteins in the Genbank database having glutamine-rich domains as defined by the submitting author.

important for dTAF<sub>II</sub>110 recognition by transcription factor activation domains. The protein encoded by A421 has such a region of alternating hydrophobic and glutamine residues, suggesting that it too could interact with dTAF<sub>II</sub>110, but of course this hypothesis would require experimental confirmation.

### 5.3.2 A421 Contains a Novel Zinc-Finger-Like Motif

Although the A421 protein has no significant homology to other proteins in the databases, a closer examination of the BlastP analysis revealed that the A421 protein contained a single copy of a zinc-finger-like motif (amino acids 446-469 of the A421 peptide sequence, figure 5.18) first identified in the *sol* gene product (Delaney *et al.*, 1991). Delaney *et al.*, (1991), considered the zinger-finger-like motif to conform to the consensus WXC<sub>2</sub>CX<sub>10</sub>CX<sub>2</sub>C (where X is any amino acid). However, I would suggest that the consensus was more accurately WXCX<sub>2</sub>CX<sub>3</sub>NX<sub>6</sub>CX<sub>2</sub>CX<sub>3</sub>R/K. An extensive search of the protein databases using the BlastP and Blitz algorithms identified a number of other proteins containing this zinc-finger-like motif, or very closely related motifs.

The proteins tabulated in figure 5.20 all contain single or multiple copies of this new zinc-finger-like motif. The additional examples of the motif indicates that the final consensus is WXCX<sub>2-4</sub>CX<sub>3</sub>NX<sub>6</sub>CX<sub>2</sub>CX<sub>3-4</sub>R/K. The spacing of the cysteine residues is similar to that found in steroid hormone receptor type zinc-fingers and the GATA-1 zinc-finger (for review see Berg and Shi, 1996), and this immediately suggests the motif represents a zinc-finger, although this has yet to be demonstrated experimentally. The question arises as to what is the function of this zinc-finger-like motif? Many of the proteins containing the zinc-finger-like motif have been the subject of functional studies, and perhaps the results of these studies may help to answer this question.

Small optic lobes (*sol*) is a *Drosophila* gene, mutations in which cause degeneration of certain classes of columnar neurons in the developing optic lobes. The gene has been cloned and sequenced (Delaney, *et al.*, 1991), and the two identified

References

Dm A421

W A C - - N M C T F R N H P Q L N I C E A C E N V - R

This chapter

Dm sol 1

W S C - - T K C N T I N P T E S L K C F N C G T V - R

Delaney et al, (1991)

2

W V C - - H A C G T D N S S V T W H C L I C D T V S Y

3

W T C - - K K C T L V N Y S T A M A C V V C G G S - K

4

W T C - - S H C T L K N S L H S P V C S A C K S H - R

5

W Q C - - P A C T Y D N C A A S V V C D I C S S P - R

HAMSTER MDM-2

Chang et al, (1995)

MOUSE MDM-2

W K C - - T S C N E M N P P L P L C N R C W T L - R

Fakharzadeh et al, (1991)

HUMAN MDM-2

W K C - - T S C N E M N P P L P S H C K R C W T L - R

Oliner et al, (1992)

HUMAN RanBP AB1

W H C - - N S C S L K N A S T A K K C V S C R N S - -

Beddow et al, (1995)

HUMAN RanBP2

W H C - - N S C S L K N A S T A K K C V S C Q N L N P

Yokoyama et al, (1995)

1

W D C - - S I C L V R N E P T V S R C I A C Q N T - K

2

W D C - - S A C L V Q N E G S S T K C A A C Q N P - R

3

W D C - - S S C L V R N E A N A T R C V A C Q N P D K

4

W D C - - S V C L V R N E A S A T K C I A C Q N P G K

5

W D C - - S V C L V R N E A S A T K C I A C Q N P G K

6

W D C - - S V C L V R N E A S A T K C I A C Q N P G K

7

W D C - - S V C C V Q N E S S S L K C V A C D A S - K

8

W N C - - N S C S F K N A A T A K K C V S C Q N T N P

MOUSE RanBP2

Wilken et al, (1995)

1

W H C - - S L C S V K N E A H A I K C V A C N N P V T

2

W E C - - S V C L V R N E R S A K K C V A C E N P G K

3

W D C - - S L C F V R N E A S A T H C I A C Q Y P N K

4

W E C - - A V C S V Q N E S S S L K C V A C E A S - K

5

W D C - - S V C L V R N E P T V S R C I A C Q N T - K

6

W D C - - S V C L V R N E P T V S R C I A C Q N T - K

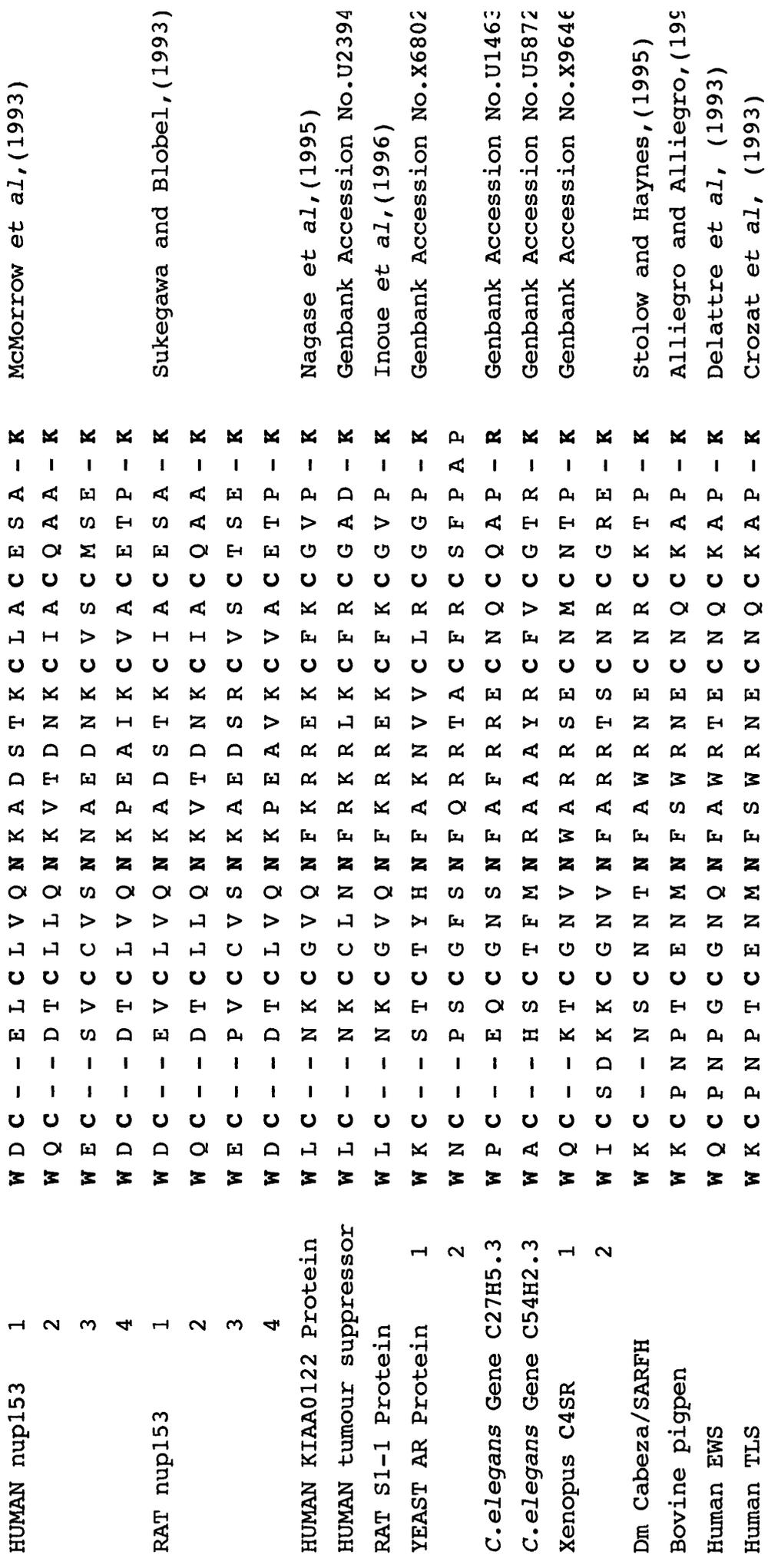


Figure 5.20. Pileup of zinc-finger-like motifs conforming to the consensus **WXCX<sub>2-4</sub>CX<sub>3</sub>NX<sub>6</sub>CX<sub>2</sub>CX<sub>3-4</sub>R/K**. Conserved residues are highlighted in bold. References are given for particular proteins where published studies are available, otherwise the Genbank accession number is given. Dm = *Drosophila melanogaster*

cDNAs, thought to represent splice variants, indicate protein products of 395 and 1597 amino acids. The larger protein product has five copies of the zinc-finger-like motif, plus a single copy of a slight variant of the form WXCX<sub>2</sub>CX<sub>4</sub>NX<sub>6</sub>CX<sub>2</sub>CX<sub>2</sub>K. I have been unable to find any other protein sequences conforming to this variant form and have therefore not included it in figure 5.18. The larger protein also includes a domain that shows similarity to the proteolytic domain of vertebrate calpains. Calpains are intracellular cysteine proteases with known substrates that include regulatory proteins such as protein kinase C (Kishimoto *et al.*, 1983). Modulation of these regulatory proteins, during the development of optic lobe neurons, is one possible requirement for sol protein. The function of the zinc-finger-like sequences in the sol proteins is unknown, but there is no suggestion that they bind DNA or nucleic acids in general.

The human nup153 and rat nup153 proteins are highly similar proteins that are integral parts of the nuclear pore complex (Sukegawa *et al.*, 1993; McMorrow *et al.*, 1994). They contain four zinc-finger-like motifs, and the region of the rat protein containing them has been expressed in *E. coli* and subject to *in vitro* DNA binding studies (Sukegawa *et al.*, 1993). The expressed protein fragment bound rat genomic DNA in a zinc dependent manner, suggesting the zinc-finger-like motifs are capable of binding DNA. Whether this is their role *in vivo* has yet to be established directly. However, the rat nup 153 protein has been located to the nucleoplasmic side of the nuclear pore complex, thereby placing it in a position to interact with genomic DNA.

The human RanBP2 and mouse RanBP2 are proteins that bind the nuclear GTPase Ran/TC4 (Yokoyama *et al.*, 1995; Wilken *et al.*, 1995). Ran/TC4 is a small G protein implicated in the initiation of DNA replication, entry and exit from mitosis, and in nuclear RNA and protein transport through the nuclear pore complex. RanBP2 is an integral part of the nuclear pore complex, and antibodies to the human protein inhibited NLS-mediated nuclear import of proteins. The human RanBP2 has eight zinc-finger-like motifs grouped together in a central region of the protein, and similarly to human nup153 this domain bound DNA *in vitro* in a zinc-dependent manner. However, an antibody to the zinc-finger-like region of human RanBP2 labelled the cytoplasmic face

of the nuclear pore complex, and similarly the murine RanBP2 was localised to the cytoplasmic rings of the nuclear pore complex. This suggests that *in vivo* the zinc-finger-like motifs of RanBP2 do not interact directly with genomic DNA.

The hamster, mouse and human MDM-2 proteins are a set of highly similar proteins implicated in tumorigenesis (Chang *et al.*, 1995; Fakharzadeh *et al.*, 1991; Oliner *et al.*, 1992). All have a single copy of the zinc-finger-like motif as well as another potential zinc-finger of the form CX<sub>2</sub>CX<sub>10</sub>CX<sub>2</sub>C. The *mdm-2* gene is over-expressed in hamster pancreatic carcinomas, human sarcomas, and in a mouse tumour cell line. MDM-2 is known to bind the sequence-specific DNA-binding protein p53 *in vitro*, and to functionally inactivate it *in vivo*. MDM-2 proteins have a nuclear-localisation signal and are thought to be restricted to the nucleus. The role of the zinc-finger-like motif in these proteins is not known.

The rat S1-1 protein is an RNA-binding protein isolated from rat liver that contains a single copy of the zinc-finger-like motif (Inoue *et al.*, 1996). The protein contains two ribonucleoprotein motifs that are thought to be the source of its RNA binding activity, and *in vitro* binding studies indicated a preference for RNA regions rich in G and U ribonucleotides. The exact function of this protein is unknown but it is thought to have a role in RNA metabolism, as other members of the S1 family of RNA-binding proteins are localised to euchromatin, and occur in association with hnRNA. Inoue *et al.*, did not recognise the existence of the zinc-finger-like motif within this protein and therefore presented no data, or made any comment, concerning its possible function. Interestingly, the zinc-finger-like motif of the rat S1-1 protein is similar to that of the putative human tumour suppressor protein, that also has RNA binding motifs, although its RNA binding activity has not been experimentally confirmed.

The *Drosophila* cabeza/SARFH, bovine pigpen, human TLS, and human EWS proteins are a class of similar proteins that also have RNA binding activity (Stolow and Haynes, 1995; Alliegro and Alliegro, 1996; Crozat *et al.*, 1993; Delattre *et al.*, 1992). Both human TLS and EWS are involved in chromosomal translocations that result in the fusion of their N-terminal coding sequences to transcription factors, resulting in

sarcoma formation. The *Drosophila* cabeza protein is also known as sarcoma-associated RNA-binding fly homologue or SARFH, and was found to be associated with transcriptionally active "puffed" regions of polytene chromosomes (Immanuel *et al.*, 1995). *Drosophila* cabeza/SARFH is a nuclear protein expressed at all developmental stages, but is particularly enriched in the CNS of late embryos, larvae and adults, and is able to bind RNA *in vitro*. Bovine pigpen is upregulated in actively migrating and proliferating endothelial cells, but barely detectable in quiescent endothelial cells. A speculative role for this class of proteins is that they bind RNA transcripts and then contact and regulate the activity of some component of the transcriptional machinery through their N-terminal domains. The function of the zinc-finger-like motif in these proteins is unknown.

Nothing is known about the function of the *Xenopus* C4SR protein (so named because of its two Cys4 zinc-finger-like motifs and its serine/arginine-rich domain), nor about the function of the Yeast AR (asparagine-rich) protein. The protein encoded by *C. elegans* gene C27H5.3 is weakly similar to nucleolins, a class of nucleolar RNA-binding proteins, but otherwise nothing is known about its function. The *C. elegans* gene C54H2.3 encodes a protein that has no significant homology to other proteins in the databases and has not been subject to functional analysis. The human KIAA0122 protein has a domain closely homologous to the RNA-binding motifs of the rat S1-1 protein, but no studies on its RNA-binding activity have been reported.

Can anything be deduced about the function of the zinc-finger-like motif from the studies outlined above? First of all it must be said that in many of these studies the existence of the zinc-finger-like motif was either not recognised at all, or only the spacing of the cysteine residues commented upon. Only McMorrow *et al.*, (1994) recognised the full extent of the zinc-finger-like motif, including the invariant tryptophan and asparagine residues and the C-terminal basic residue, and only Sukegawa *et al.* (1993), and Yokoyama *et al.* (1995) subjected the zinc-finger-like motifs to explicit functional analysis. In both of these latter cases they found the protein domains, containing the zinc-finger-like motifs, bind genomic DNA in a zinc-

dependent manner. These proteins are both integral members of the nuclear pore complex but are located on the nucleoplasmic and cytoplasmic sides respectively. This casts doubt on whether the zinc-finger-like motifs bind genomic DNA *in vivo*, as a protein located on the cytoplasmic side of the nuclear pore complex would presumably not be in a position to do so. However, Wilken *et al.*, (1995) speculate that the cytoplasmically located zinc-finger-like motifs may protect the nucleus from foreign DNA molecules (e.g. viral DNA) by their binding and immobilisation.

There are other examples of zinc-fingers that bind DNA *in vitro* but where they are not thought to play this role *in vivo*. For instance, sequence analysis of the Really Interesting New Gene 1 or *ring1* led to the identification of a novel zinc-finger of the C<sub>3</sub>HC<sub>4</sub> type which became known as the RING-finger motif (Freemont *et al.*, 1991). The RING-finger motif can be described as CX<sub>2</sub>CX<sub>9-39</sub>CX<sub>1-3</sub>HX<sub>2-3</sub>CX<sub>2</sub>CX<sub>4-48</sub>CX<sub>2</sub>C where X can be any amino acid, although there are clear preferences for particular residues at certain positions. There is evidence that RING-containing proteins can bind DNA in a zinc-dependent manner *in vitro* (Lovering *et al.*, 1993). However, the RING-domain is found in cytoplasmic and integral membrane proteins, and it is increasingly clear that RING-finger-containing proteins are often part of multi-protein complexes (Borden and Freemont, 1996). It has been suggested that the RING-finger mediates the formation of these multi-protein complexes, and indeed that the RING-domain acts directly as some type of 'glue' in forming protein-protein assemblages (Saurin *et al.*, 1996).

The LIM domain is another cysteine-rich zinc-finger motif, first identified in *Caenorhabditis elegans* (Way and Chalfie, 1988), that is thought to play a role in protein-protein interactions. Sequence analysis of LIM proteins produces the following consensus for the LIM domain, CX<sub>2</sub>CX<sub>17-19</sub>HX<sub>2</sub>CX<sub>2</sub>CX<sub>2</sub>CX<sub>16-20</sub>CX<sub>2-3</sub>C, although the final cysteine can be histidine or occasionally aspartic acid. Proteins containing LIM domains fall into two classes, the first class has two LIM domains associated with a DNA-binding homeodomain (designated LIM-HD proteins), and the second class of proteins have only one to three LIM domains (designated LIM-only proteins). LIM-HD

proteins have been implicated in the control of cellular differentiation, for example, the *lin-11* gene encodes a protein that controls certain asymmetric cell divisions during vulval development in *C. elegans* (Way and Chalfie, 1988). Similar instances where LIM-HD proteins are involved in the control of development in other organisms are known, for example in *Xenopus* (Taira *et al.*, 1992) and *Drosophila* (Bourgouin *et al.*, 1992). LIM-only proteins appear to have a diverse range of functions from the pollen-specific plant protein SF3 (Baltz *et al.*, 1992), to the serum-response protein CRP (Wang *et al.*, 1992). Some of these LIM-only proteins are localised in the cytoplasm, and it has been suggested that the LIM domain is involved in protein-protein interactions (Sanchez-Garcia and Rabbitts, 1994).

Other functions for zinc-fingers are not unknown. For instance, protein kinase C contains two zinc-fingers, of the form CX<sub>2</sub>CX<sub>13</sub>CX<sub>2</sub>CX<sub>7</sub>CX<sub>7</sub>C, that are known to bind the regulatory molecules phorbol ester or diacyl glycerol, rather than mediate protein-protein interactions or bind DNA (Ono *et al.*, 1989).

I would suggest that based on the functional diversity of the proteins containing the zinc-finger-like motif that it is unlikely to be involved directly in DNA or RNA binding. The RNA-binding proteins described earlier all contain distinct RNA-binding domains, that have, in some cases, been shown to be both necessary and sufficient for RNA binding. It therefore seems unlikely that the zinc-finger-like motif would also play this role. Furthermore, in the A421 protein the zinc-finger-like motif is immediately next to a highly acidic region (see figure 5.18), which would probably prevent this region of the protein interacting with negatively-charged nucleic acids. It is clear from the studies on the zinc-finger-like-motif-containing proteins that protein-protein interactions are important to the function of many of them, and it is obviously a possibility that the zinc-finger-like motif mediates these protein-protein interactions in a similar way to the RING and LIM zinc-fingers. Alternatively the zinc-finger-like motif may bind some regulatory macromolecule, in an analogous manner to the zinc-fingers of protein kinase C.

Assuming that the zinc-finger-like motif is not involved directly in nucleic acid binding, it is remarkable that many of the proteins containing the zinc-finger-like motif, for which functional studies have been carried out, are clearly nuclear proteins; either integral constituents of the nuclear pore complex (nup153 and ranBP2), RNA-binding proteins involved in transcriptional regulation (cabeza etc.), proteins that interact with the DNA-binding protein p53 (MDM-2), or RNA-binding proteins involved in RNA metabolism (rat S1-1). The one exception appears to be the protein products of the *Drosophila sol* gene, although the exact cellular localisation of these is not known, and thus a nuclear function for these proteins can not be ruled out. Interestingly, the A421 protein has a putative nuclear localisation signal conforming to the rules worked out by Boulikas, (1994), i.e. a hexapeptide containing four basic (arginine or lysine) residues, but without acidic (glutamic acid or aspartic acid) or bulky (phenylalanine, tyrosine or tryptophan) residues (figure 5.18). Thus it is entirely possible that the A421 protein is also restricted to the nucleus.

The zinc-finger-like motif is an unusual example of zinc-fingers in having invariant tryptophan and asparagine residues, together with the usual presence of a C-terminal basic residue. Although zinc-fingers motifs are usually defined by the spacing of cysteine or cysteine and histidine residues, in practice other residues are often conserved. The TFIIA and LIM type zinc-fingers both contain conserved hydrophobic residues that form a hydrophobic pocket stabilising the tertiary structure of the zinc-finger. However, it is unlikely that the polar asparagine and basic residues play a similar role in the zinc-finger-like motif. Perhaps they stabilise the zinc-finger in some other way, or play a more direct role in the functioning of the zinc-finger. But these questions will only be answered by experimentation, either by introducing mutations into the zinc-finger-like motif, or by more direct structural studies using X-ray crystallography or NMR.

### 5.3.3 Does the A421 gene product contain a Leucine Zipper?

If the glutamine-rich region of A421 represents the activation domain of a transcription factor, it might also be reasonably expected to be a DNA-binding protein. As discussed earlier I think there are serious reservations about whether the zinc-finger-like motif fulfills a DNA-binding role. Furthermore, as also noted earlier, the A421 protein has no significant homology to other proteins in the data bases, nor has it any of the recognised protein motifs in the Prosite database, indicating it has no obvious DNA-binding domain. However, a closer examination of the protein sequence reveals a potential solution to the problem. The region from amino acid 211 to 238 of the A421 protein sequence is particularly rich in basic residues, and is followed by a region that has a number of leucine residues (figure 5.18). This is highly suggestive of a possible leucine zipper motif. The leucine zipper is a well characterised DNA-binding motif, consisting of a basic domain followed by four or five leucines spaced every seventh residue (reviewed in Busch and Sassone-Corsi, 1990). Leucine zippers have an  $\alpha$ -helical structure and form dimers, in a coiled-coil formation, to produce a functional complex that binds DNA via the basic domains. The heptad spacing of the leucines is critical to dimer formation, as in the coiled-coil structure the helical repeat is 3.5 residues per turn, placing all the leucines on the same side of the  $\alpha$ -helix. It is the hydrophobic interaction between leucines, from interfacing parallel  $\alpha$ -helices, that is the basis of dimer formation.

An alignment between the A421 protein sequence, from amino acid 210 to 275, and some known leucine zipper domains is shown in figure 5.21. As can be seen, the A421 sequence has many features typical of leucine zippers, in particular, two basic regions separated by a spacer containing alanines, followed by a region containing leucines (including one isoleucine) spaced every seventh residue. However, the alignment between the last three leucines of the A421 putative leucine zipper, and the leucines of the other leucine zippers, is only achieved by omitting residues 249 to 259 of the A421 sequence. Without this omission the final three leucines are not part of the

GCN4	SSDPAAL <u>KRARNT</u> EAARRSRARKLORMKQLEDKVEELLSSKNYHLENEVARLKKLVGER
YAP-1	DPE <u>TQKR</u> RTAQNRAAQRAF <u>REKERKER</u> KKVQSLESIQQQNEVEATFLRDQLITL
C/EBP	RRERNNIKV <u>RRSRD</u> KAQORNVE <u>TQQK</u> VLELTSDNDRLRKRRL <u>EQLSRE</u> LDTLRGIFRQL
CREB	AARKREVR <u>LMKNRE</u> AAARECRKRKKEYVKCLENRVAVLENQNKTLIEELKALKDLYCHK
CRE-BP1	PDEKRRKFLERNRAAA <u>SR</u> CROKRKVVQ <u>SLEKKAED</u> SSLNGQLQSEVTLLRNEVAQL
c-JUN	ERIKAE <u>RKRMRN</u> RIAAS <u>KCRK</u> RKLERIARLEEKVKTLKAQNSELASTANMLREQVAQL
c-FOS	EE <u>KRRIR</u> RRERNKMAAAKCRNRRREL <u>TD</u> TLOAETDQLEDEKSALQTEIANLLKEKEKL
A421	DE <u>RR</u> AA <u>TIER</u> QKORRDKLANALRDNKKRLLVLEQEINI~~~RLDRDIKQLTEDCNRL

Figure 5.21. Alignment between A421 sequence (amino acids 210-275) and the sequences of some known leucine zippers (Adapted from Busch and Sassone-Corsi, 1990). Basic regions are double underlined. Alanines of spacer region are underlined. Leucines are highlighted in bold. Amino acids 249-259 of A421 sequence not included indicated by ~~~.

same heptad repeat pattern as the earlier leucines. Thus, if the entire region was  $\alpha$ -helical in structure, all the leucines would not occur on the same side of the helix, as required by the model for leucine zipper dimer formation. However, an examination of the predicted secondary structure of the A421 protein suggests a possible solution (see figure 5.22) The region containing the putative leucine zipper is generally predicted to have the necessary  $\alpha$ -helical structure. But crucially, a break occurs in the  $\alpha$ -helix, in the region of amino acids 249 to 257, because of the presence of two prolines. Perhaps then, the A421 protein contains a leucine zipper of a new type, with two  $\alpha$ -helical regions each with leucines that participate in hydrophobic interactions. The first helical region would contain the basic domain and two leucines and an isoleucine, with a loop of undefined structure separating it from the second helical region, which contains the other three leucines. As depicted schematically in figure 5.23, it might therefore be possible for all the leucines and the isoleucines, on the two separate helical regions, to participate in the hydrophobic interactions necessary for dimer formation.

#### 5.3.4 Conclusion

What conclusions can be drawn about the A421 gene? From the reverse Northern and Northern analysis this gene has a transcript that clearly has head elevated expression, and the *in situ* hybridisation data suggests this head expression is within neurons of the *Drosophila* brain. However, it would appear that this neuronal expression is not restricted to the mushroom bodies, and therefore does not reflect expression of the 201Y P[GAL4] enhancer-trap. Analysis of the A421 protein sequence reveals a glutamine-rich region which is suggestive of a transcriptional activation function. However, many proteins with glutamine-rich regions are not transcription factors and further experimental work would be necessary to establish this conclusively. The A421 protein has a novel zinc-finger motif, which I would suggest is involved in protein-protein interactions rather than nucleic acid binding. It also contains a potential leucine-zipper of non-standard structure that may be involved in DNA binding, but again,



```

A421 sequence
Gibrat method
Levin method
DPM method
SOPMA method
PhD method
Consensus

321 | 330 | 340 | 350 | 360 | 370 | 380 | 390 | 400 |
|  |  |  |  |  |  |  |  |  |
RQARVQAPPSSRLHVSVAAPNSQPVQDFGQHHSSAPTSACLTPMQSQQLLOLQEQPPPTAQQYQFQQYLOQQRQQ
HH-EEE---HHEEE---EEEEEE---EEE---HHHHHHHH---HHEHEHEHHHHHHHHHHHH
-E---S-EEE---S---S---HEEEES---EE-THHHHHHHH-H---H
HHHHHH---HHEEEE-H-T---E-E-TT---HEE---HHHHHHHHHH---HHEEHEHEHEHHHHHHHH
-----EEEE-----HHHHHHHHHHHHHH---HHHHHHHHHHHH---HHHHHHHHHHHH
-----EEEE-----HHHHH-----H-----HHHHHHHHHHHH

401 | 410 | 420 | 430 | 440 | 450 | 460 | 470 | 480 |
|  |  |  |  |  |  |  |  |  |
QLQQMQQQQLQQMQQQQPQQPQQEFLSDSDVDEEEETLDSWACNMCNFRNHPQLNICEACENVRIQPGMIRIVPS
HHHHHHHHHHHHHHHHHH---HHHH---HHHHH-HHHH---E---HHHH---H-H-----EEEE--
EHH-HHH-HHH-HH-S---HHHHHHHT---HHHHHHHHHHH-T-S-T---HHHH---TS-EEEE--
HHHHHHHHHHHHHHHHHH---HHHHHT---TEHHHHHHHH-HHHHEEE---E-EHHHHHHHEE---EEEE-T
HHHHHHHH---
HHH-HHHHHH---
HHHHHHHHHHH-HH-HH---HHHH---HHHH---HHH-----H-----EE--
HHHHHHHHHHHHHHHHHHHHHHHHHH---HHHHHH---HHH-----HH-----EEEE--

481 | 490 | 500 | 510 |
|  |  |  |  |
GGGAAAAATPPGSIEQQQPPSQPYALHT
---HEEE---EEE-----HHEEEE
---HHH-----H---HEETT
T---HHHH---EHHHT-----HH--
-----
-----EE--
-----EE--
-----EE--

```

Figure 5.22. Predicted secondary structure of A421 protein, using the Protein Sequence Analysis package (Institute of Biology and Chemistry of Proteins, Paris. <http://www.ibcp.fr/predict.html>). H =  $\alpha$ -helix, E = extended ( $\beta$ -sheet), S = Bends, T = Turn. Gibrat method, Gibrat *et al.*, (1987); Levin method, Levin *et al.*, (1986); DPM method, Deleage and Roux, (1987); PhD method, Rost and Sander, (1994); SOPMA method, Geourjon and Deleage, (1994).

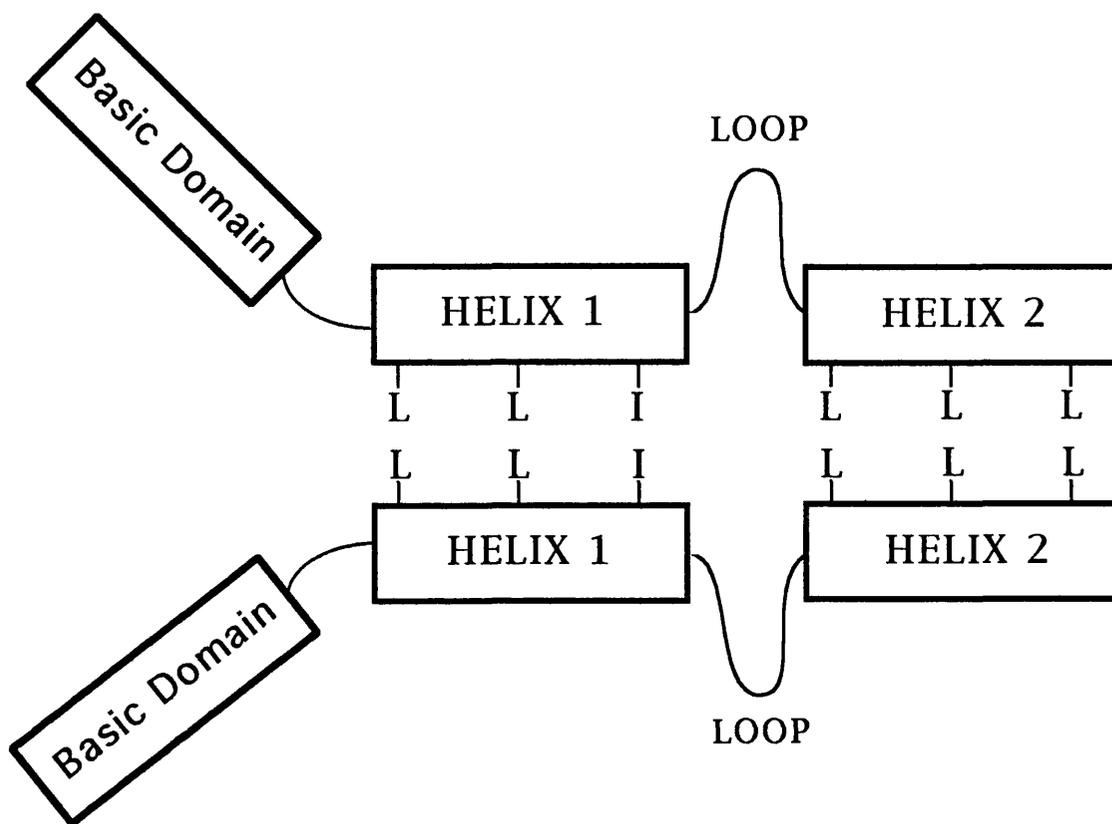


Figure 5.23. Schematic representation of the possible structure of the A421 leucine zipper. Potential interaction between leucines on two separate helical regions might allow dimer formation.

further experimental work would be necessary to establish this. I think a not unreasonable working hypothesis is that the A421 gene encodes a novel transcription factor that has a particular important function in *Drosophila* neurons.

**Chapter 6**  
**Generation of 201Y Excision Mutants**

## 6.1 Introduction

The fly line 201Y is viable when homozygous for the P[GAL4] insertion, and analysis presented in earlier Chapters suggests that the 201Y P[GAL4] insertion does not disrupt expression of the adjacent genes. As part of the analysis of the function of these adjacent genes it would be useful to have mutant or null alleles. P-elements can, in the presence of transposase, be remobilised leading to a number of different kinds of excision events, i.e. internal excisions, precise excisions, and imprecise excisions (figure 6.1). Imprecise excisions, removing the P-element and flanking genomic DNA, can generate new alleles of adjacent genes. Deletion of flanking sequences is usually one sided (i.e. imprecise excisions 1 and 2 of figure 6.1), and can range from a few bp to many kb in length. New alleles of *rudimentary* and *Sex-lethal* have been produced in this way (Tsubota and Schedl, 1986; Salz *et al.*, 1987). As the A421 and FKBP12 genes appear to be single-copy, imprecise excisions that remove these genes, would also totally remove their gene products, and allow the assessment of any mutant phenotypes.

## 6.2 Generation and analysis of excision lines

The genetic scheme used to generate excised P-elements is outlined in figure 6.2. The logic of the scheme is basically that flies homozygous for the 201Y P[GAL4] insertion are crossed to a source of transposase contained on the  $\Delta 2-3$  P-element. Flies containing the 201Y P[GAL4] and  $\Delta 2-3$  P-elements are then crossed to flies carrying genetically marked second chromosomes, and white eyed progeny (i.e. an excision event has occurred removing the *white*<sup>+</sup> gene carried by the P[GAL4] element) not carrying the  $\Delta 2-3$  P-element are selected. These are individually crossed to flies carrying genetically marked second chromosomes, and the progeny crossed to produce pure breeding lines.

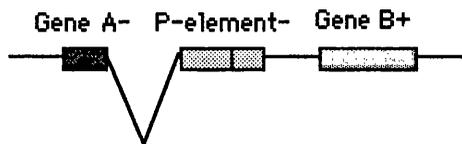
**Pre-mobilisation situation**



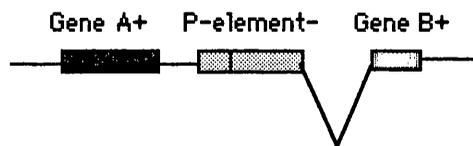
**Precise Excision.**



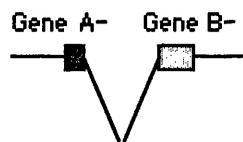
**Imprecise Excision - 1**



**Imprecise Excision - 2**



**Imprecise Excision - 3**



**Internal Excision**



**Figure 6.1.** Potential excision events following P-element mobilisation from a hypothetical locus. Imprecise excision can lead to new alleles of adjacent genes.

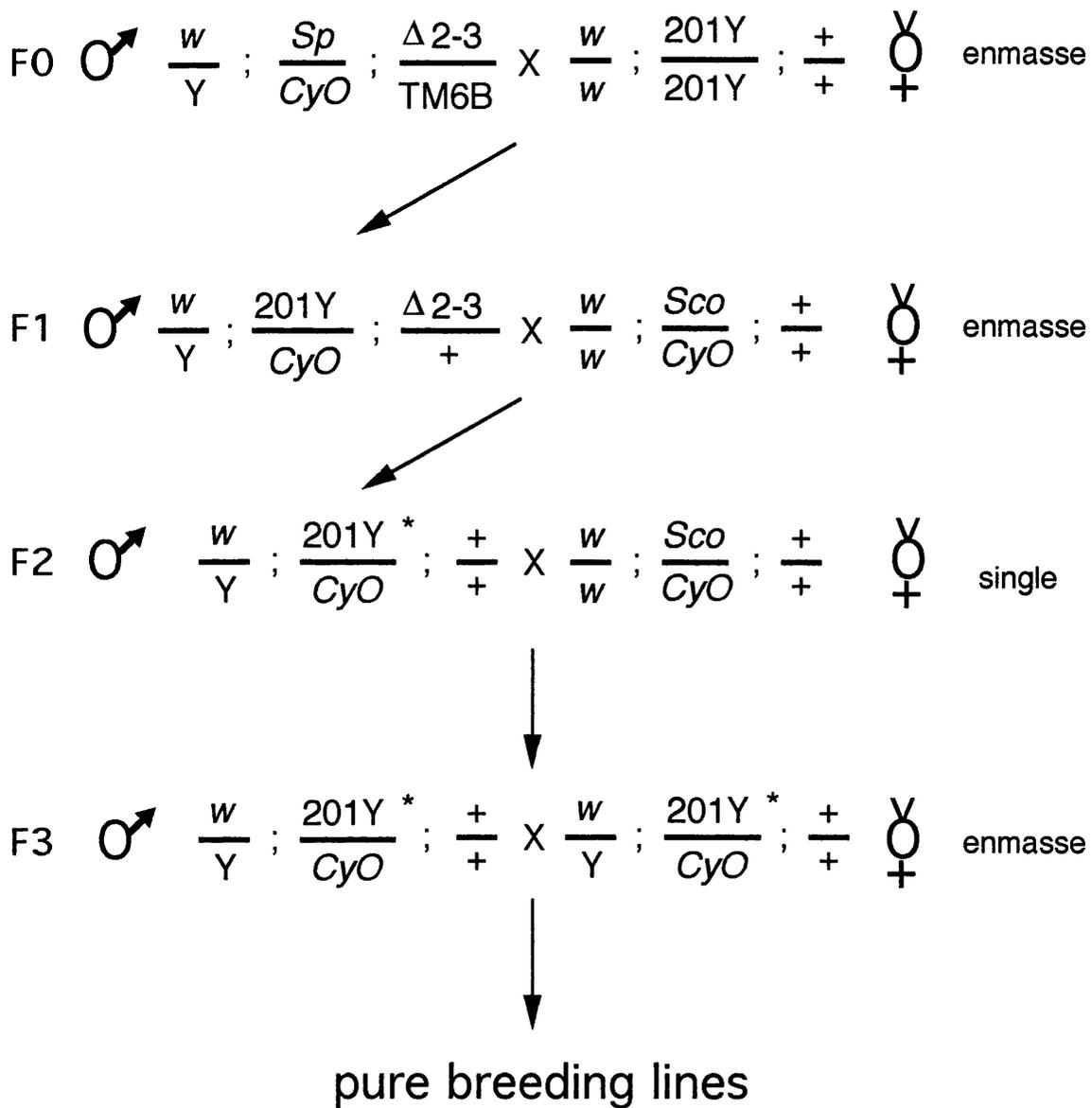


Figure 6.2. Generation of excision strains via P-element mobilisation.  $\Delta 2-3$  refers to the *Dr* P[ry<sup>+</sup>  $\Delta 2-3$ ]99B chromosome. 201Y refers to chromosomes containing the normal 201Y P[GAL4] element. 201Y\* refers to chromosomes in which an excision event has occurred.

Approximately 4000 F2 flies were screened and 36 white eyed flies isolated. From the 36 fly lines established, fly lines 3, 18, 19, 28, and 36 appeared to carry lethal mutations on the second chromosome, i.e. these fly lines produced only *Cy* progeny. All of the other excision lines appeared perfectly viable and had no obvious mutant phenotype. To check if the lethal mutations occur in the same gene or represent a single complementation group, reciprocal crosses between the different lethal strains were performed. Results of these crosses are shown in figure 6.3., and suggest that the lethal fly lines fall into three complementation groups. One complementation group consists of fly lines 3, 19, and 28, with the other complementation groups consisting of the single fly lines 18 and 36.

To help determine if the lethal mutations represent mutations at cytological location 56D, i.e. the location of the 201Y P[GAL4] insertion, fly line Tp(2;Y)jl29 was acquired (gift from Prof. T. W. Lyttle, University of Hawaii. (Lindsley and Zimm, 1992)). This fly line carries two types of second chromosome, one that is deficient for 54A-56F, and the other that is *cn bw* but otherwise structurally normal. The missing material from the deficient second chromosome is inserted on the Y chromosome. The strain is maintained by crossing, at each generation, males with wild-type eye colour (Tp(2;Y); Df/*cn bw* or Tp(2;Y); Df/Df) to white-eyed females (*cn bw/cn bw*). This strain can be used to check the lethal mutations generated in the P-element excision in two independent ways. First, flies heterozygous for the lethal mutation and the deficiency should be lethal if the mutation is at 56D. Second, the transposed material on the Y chromosome should allow male flies homozygous for the lethal mutation to be viable, if the mutation is at 56D. The genetic schemes used to test both of these possibilities are shown in figures 6.4 and 6.5.

The results of these genetic analyses are that only fly lines 3, 19 and 28 carry lethal mutations that are not viable when heterozygous with the 54A-56F deficiency, and produce viable TpY males that are homozygous for the lethal mutation. In contrast, fly lines 18 and 36 carry lethal mutations that were viable when heterozygous with the 54A-56F deficiency, but did not produce viable TpY males that were homozygous for

♀ ♂	3	18	19	28	36
3		+	-	-	+
18	+		+	+	+
19	-	+		-	+
28	-	+	-		+
36	+	+	+	+	

Figure 6.3. Grid indicating results of crosses between lethals isolated during generation of excision lines. + = complementation, i.e. heterozygotes were viable. - = no complementation, i.e. heterozygotes were lethal.

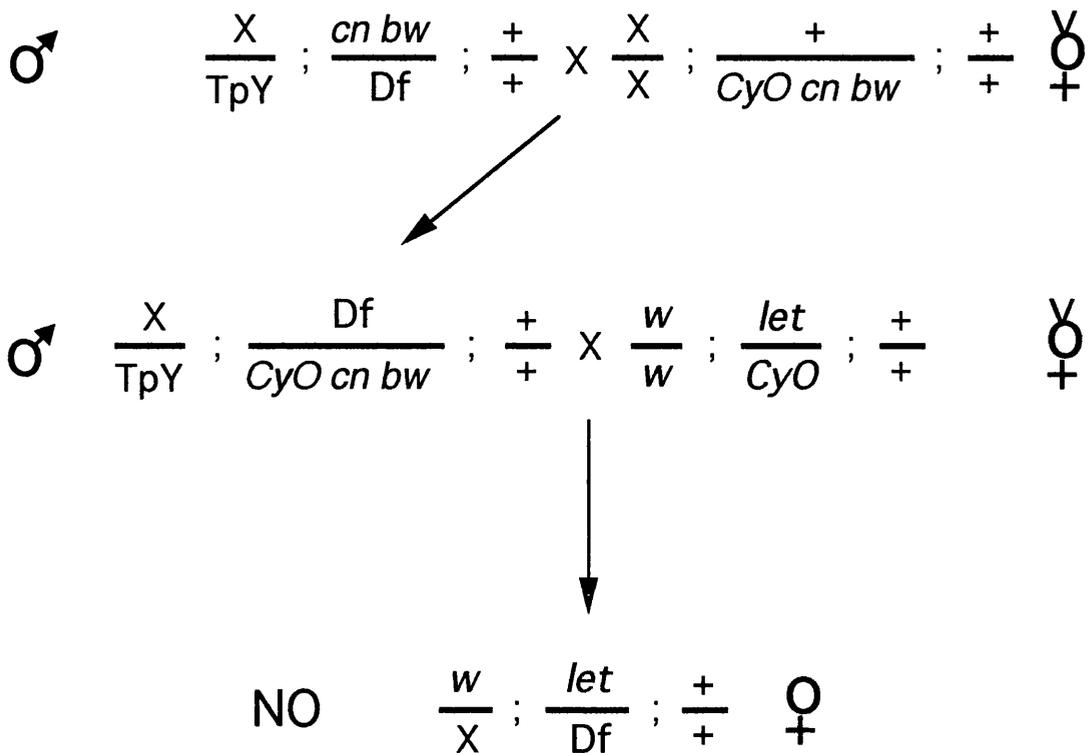


Figure 6.4. Deficiency analysis of lethal strains. *let* refers to chromosomes carrying lethal mutations isolated during generation of excision lines. *Df* refers to a chromosome lacking the 54A-56F region. *TpY* refers to a Y chromosome carrying the 54A-56F transposed material. The initial cross is to generate flies carrying the *Df* and a second chromosome with a dominant genetic marker. If the lethal mutation is in the 54A-56F region, the final cross should generate only *Cy* females.

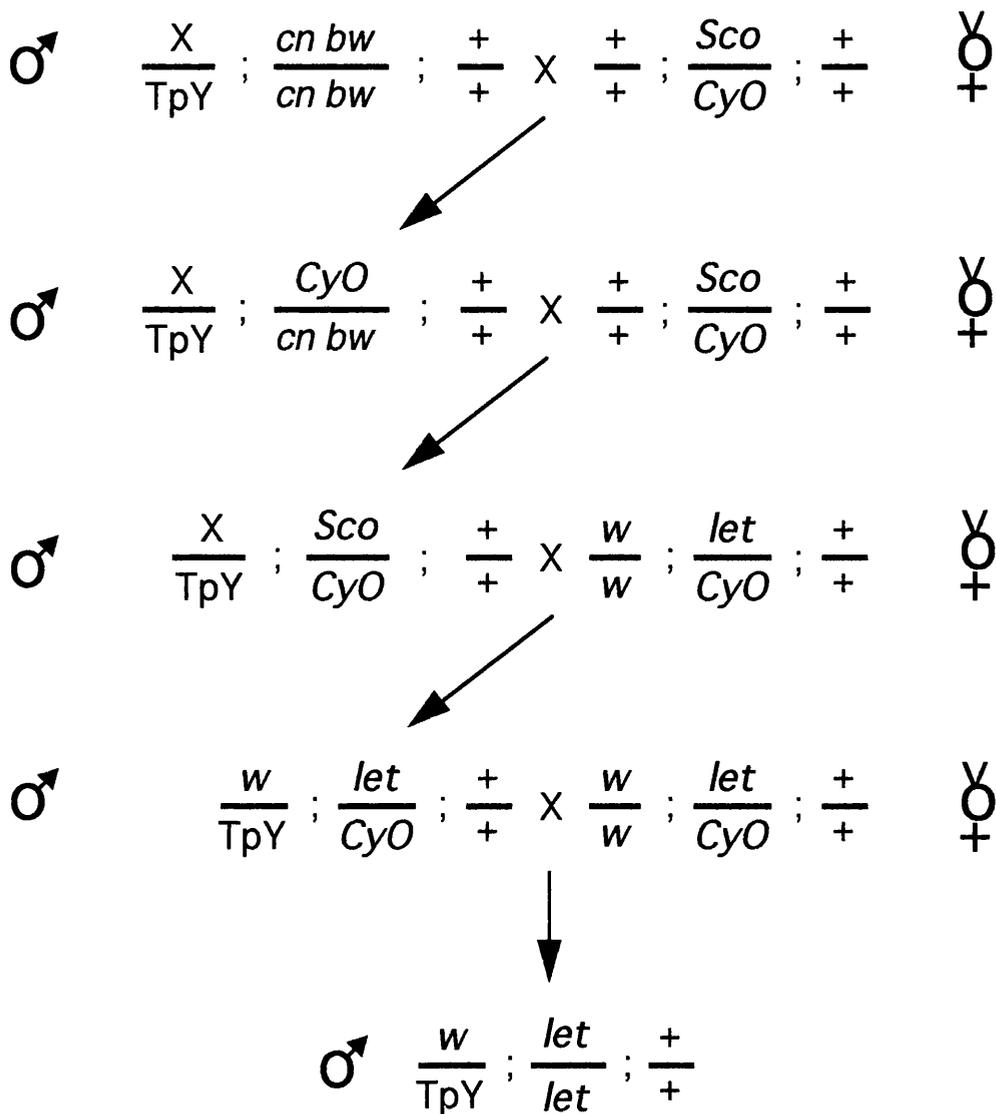


Figure 6.5. Analysis of lethals with TpY. First two crosses are to generate flies carrying the TpY chromosome, and second chromosomes with dominant genetic markers. *let* refers to chromosomes carrying lethal mutations isolated during generation of excision lines. TpY refers to a Y chromosome carrying the 54A-56F transposed material. If the lethal mutation is in the 54A-56F region, the final cross should generate non-Cy males.

the mutation. Presumably these two fly lines carry mutations that are not within the 54A-56F region, and are therefore not mutations caused by imprecise excision of the 201Y P[GAL4] element. They may represent spontaneous mutations that had accumulated in the 201Y strain, or alternatively, they may be mutations caused by insertion of the P-element at a new location.

As discussed earlier fly lines 3, 19, and 28 are members of the same complementation group, and this further genetic analysis adds weight to the thesis that they represent mutations in the same gene. Of course this genetic analysis does not indicate which, if either, of the two genes flanking the 201Y P[GAL4] element has been disrupted. Therefore to determine if the FKBP12 or the A421 gene have sequences deleted by imprecise excision of the P[GAL4] element requires molecular analysis. Genomic Southern blots were performed on DNA from each of the 36 fly lines. Figure 6.6 shows the result of probing approximately 5 $\mu$ g of genomic DNA digested with *XhoI*, with the 2.4kb *XhoI* fragment of pB201Y labelled with <sup>32</sup>P. For all fly lines a single wild-type band of 2.4 kb is evident. The DNA fragment used as a probe contains the FKBP12 gene analysed in Chapter 4, therefore the genomic Southern result suggests that no disruption of this gene has occurred in any of the excision fly lines. However, the lethal fly lines were necessarily analysed as heterozygotes, with the balancer chromosome having a wild-type 201Y genomic locus giving a wild-type band in the genomic Southern. Another band indicating the presence of a deletion, would only be visible if the deletion break-point occurred within the region covered by the probe. If the deletion removes the entire region covered by the probe then only the wild-type band provided by the balancer chromosome will be evident. Figure 6.7 shows the result of probing *SstI* digested genomic DNA with the 5.5kb *SstI* fragment from genomic clone 2.1 labelled with <sup>32</sup>P. This genomic fragment contains the A421 gene. Only the wild-type band is evident again. However, the same caveat relating to the lethal lines, as applied to figure 6.6, is applicable here. Clearly a more sophisticated molecular analysis of the lethal fly lines is required.

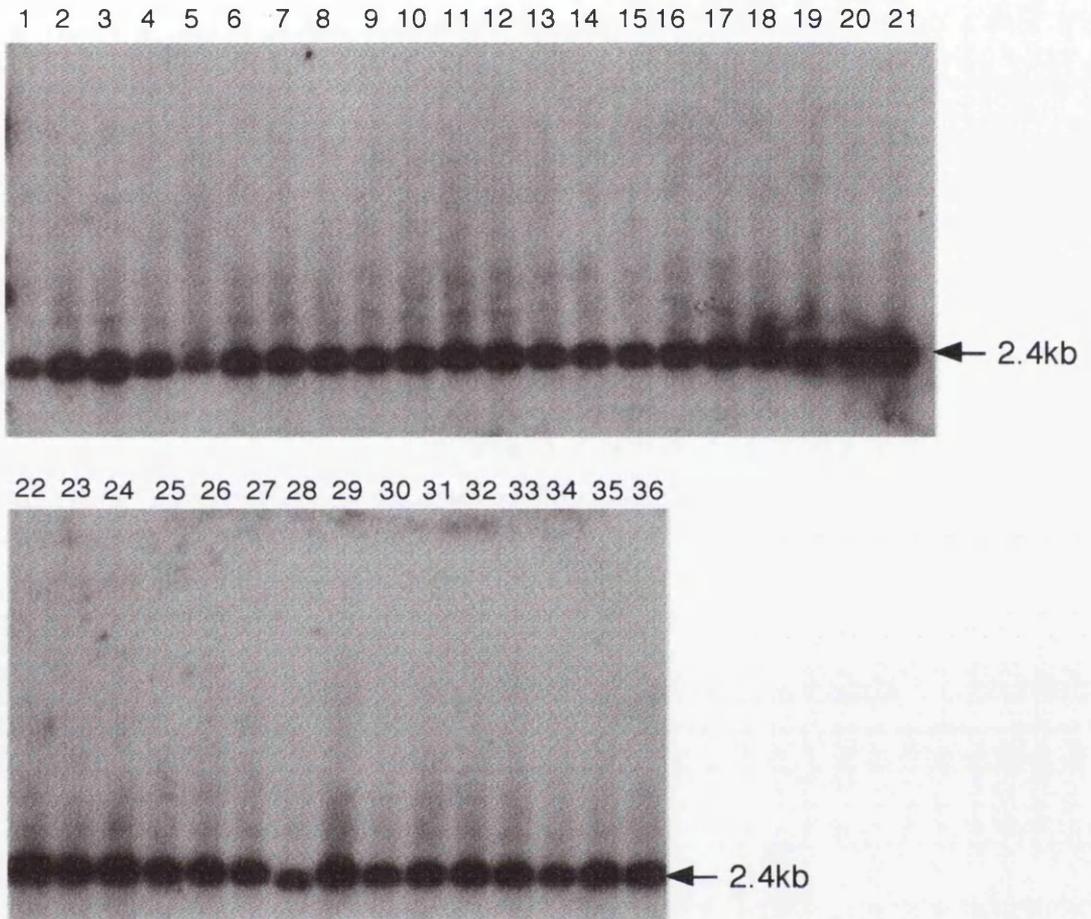


Figure 6.6 Southern blot analysis of excision lines. Genomic DNA was digested with *XhoI*, blotted and probed with the 2.4kb *XhoI* fragment from pB2101Y. In all cases a wild-type 2.4kb band is evident. Lane numbers refer to excision lines.

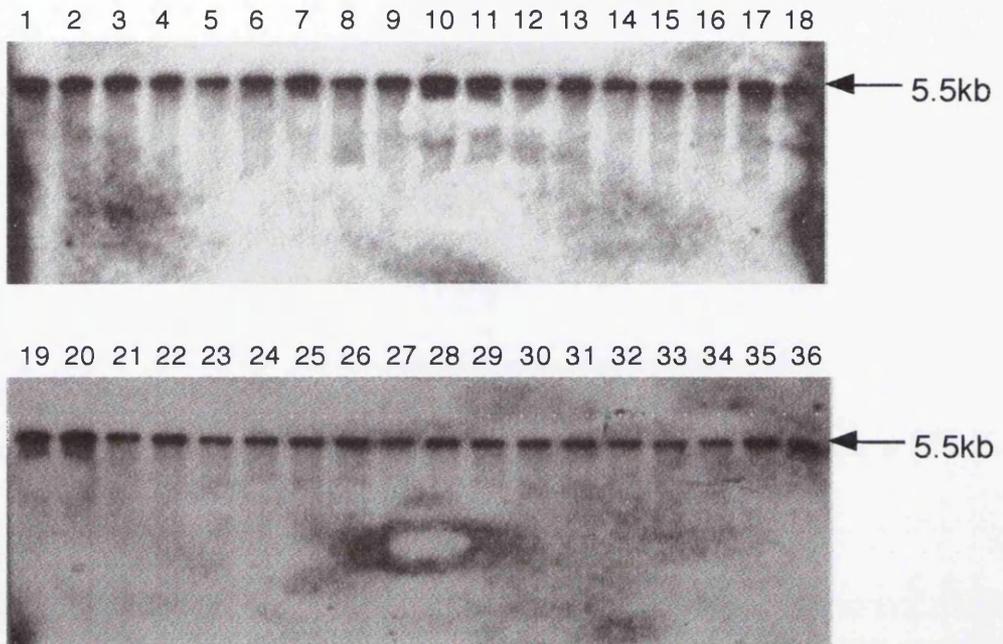


Figure 6.7. Southern blot analysis of excision lines. Genomic DNA was digested with *SstI*, blotted and probed with the 5.5kb *SstI* fragment of genomic clone 2.1. In all cases a wild-type 5.5kb band is evident. Lane numbers refer to excision lines.

### 6.3 Discussion

Genetic analysis of the 201Y P[GAL4] excision lines suggests that three lines, 3, 19, and 28, carry lethal mutations that comprise a single complementation group. No known lethals map to cytological location 56D so these three lines potentially represent a new lethal mutation. However, analysis of these lines by genomic Southern blots gave inconclusive results, and it is therefore not possible to say which, if either, of the two genes flanking the 201Y P[GAL4] insertion has been disrupted. However, it is known that mutations in the yeast FKBP12 gene are not lethal (Heitman, 1991). Furthermore, the family of FK506 binding proteins has a diverse array of members, and it is probable that most eukaryotes have a number of different FKBP proteins of related and overlapping function. Therefore, disruption of any one FKBP gene may not have a serious deleterious effect. Within *Drosophila* two members of the FKBP family of proteins are known, FKBP12 and FKBP39, and it is probable that others remain to be discovered. I would therefore suggest that disruption of the *Drosophila* FKBP12 gene would not give a lethal phenotype.

To determine if disruption of the A421 gene is the cause of the lethal phenotypes requires a more sophisticated analysis than the simple genomic Southern blots discussed in section 6.2. Further genomic Southern blots using probes comprised of other genomic fragments may uncover the excision breakpoints. Genomic Southern blots performed in a quantitative manner may determine whether the heterozygotes carry only one copy of the A421 gene. If oligonucleotide primers spanning the imprecise excision could be designed then PCR could determine if the A421 gene had been disrupted or lost. However, this would require some idea of the extent of the imprecise excision, and the sequencing of genomic clones to design primers.

**Chapter 7**  
**Genes Flanking the 43Y Enhancer-Trap**

## 7.1 Introduction

43Y is a *Drosophila* line containing a single P[GAL4] enhancer trap, inserted at cytological location 2C (Zonsheng Wang, pers. comm.). Line 43Y exhibits enhancer-trap expression in the mushroom body intrinsic neurons, Kenyon cells, and mushroom body extrinsic neurons. 43Y also has enhancer-trap expression in a small cluster of cell bodies located at the frontal dorsal margin of the brain, from which processes arise to terminate within the  $\gamma$ -lobe, and within the lateral protocerebrum (Armstrong, 1995). Plasmid-rescue had generated both 3' and 5' plasmids for this fly line, both of which contain transcriptional units as revealed by reverse Northern analysis. This chapter describes identification and analysis of the genes identified by reverse Northern analysis. (Some of the work described in this chapter was performed by a project student, Liam Nielson. The subcloning and sequencing of cDNAs, and the immunohistochemistry were largely his work.)

## 7.2 Analysis of the 43Y 3' gene

### 7.2.1 Restriction mapping of pB43Y3'

Plasmid rescue of 3' flanking genomic DNA, digested with restriction enzyme *SstI*, yielded a plasmid containing 1.8kb of genomic DNA, upon which the reverse Northern analysis described in Chapter 3 was performed. Further work, using restriction enzyme *PstI*, generated another larger plasmid containing 3.2kb of 3' flanking DNA. Restriction digests of this plasmid with *Sall*, *SstI* and *XhoI* gives fragment sizes of 3.5kb and 2.7kb, 4.9kb and 1.3kb, and 4.95kb and 1.25kb respectively (figure 7.1). To assist with the production of an accurate restriction map for this plasmid double digests with the above enzymes and *BamHI* (an enzyme that cuts in the extreme 3' polylinker) were performed. Digestion with *Sall* and *BamHI* gives fragments of 3.3kb, 2.5kb and 0.4kb, with *SstI* and *BamHI* gives fragments of 3.0kb, 1.9kb, 0.9kb

and 0.45kb, with *BamHI* and *XhoI* giving fragments of 3.0kb, 1.9kb, 0.95kb and 0.4kb. From analysis of these fragments the restriction map depicted in figure 7.2. can be constructed. This plasmid is clearly related to the smaller *SstI* generated plasmid, used in the reverse Northern analysis, as it contains a *SstI* site approximately 1.8kb from the 3' end of the P[GAL4] element.

### 7.2.2 Isolation and sequence analysis of cDNA clones

To facilitate identification of genes contained within the genomic DNA of rescued plasmid pB43Y3', a cDNA library was screened. An adult *Drosophila* head cDNA library, constructed in lambda phage NM1149, was screened, using the 2.9kb *BamHI* genomic DNA fragment of pB43Y3' labelled with <sup>32</sup>P as a probe. The primary screen was conducted on 100 000 recombinant phage plaques. Ten positive phage were identified and purified down to single plaques. Phage DNA was prepared from each positive and digested with restriction enzymes *EcoRI* and *HindIII* to release the cDNA insert. The DNA fragments were separate by gel electrophoresis, Southern blotted onto nylon filters, and probed with the <sup>32</sup>P labelled 2.9kb *BamHI* fragment of pB43Y3' to confirm them as real positives (Figure 7.3). Nine of the positive phage gave identical cDNA fragments of 1.2kb in size, and all of these nine fragments hybridised to the pB43Y3' probe, confirming them as real positives. The tenth positive gave a cDNA fragment of 1.6kb that did not hybridise to the pB43Y3' probe, therefore further analysis of this cDNA was not pursued.

The positive 1.2 kb cDNA phage fragment was cloned into the plasmid pBluescript (creating a recombinant plasmid known as pB43c1) to allow the rapid acquisition of sequence data. Using T3 and T7 primers, which hybridise to specific sequences either side of the multiple cloning site of pBluescript, sequence data from the 5' and 3' ends of the cDNA were obtained. These data were used to search DNA sequence databases using the BlastN algorithm. Both sequences are almost identical to sequences from the *ultraspiracle (usp)* gene of *Drosophila*. The alignments between

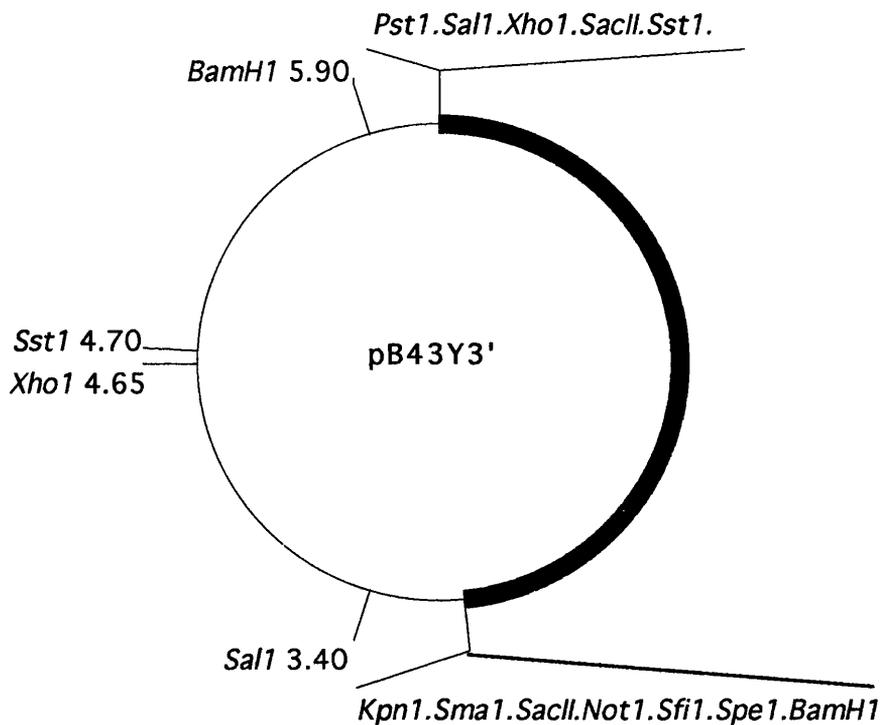


Figure 7.2 Restriction map of pB43Y3'. Numbers indicate distance in kb from "twelve o'clock" position. Black bar represents *amp<sup>R</sup>* and origin of replication from the P[GAL4] element.

sequences from pB43c1 and *usp* are shown in figure 7.4. The few misalignments are assumed to be errors in the data from the single sequencing runs from pB43c1. The pB43c1 cDNA sequences span a region from nucleotide 1138 to nucleotide 2294 of the *usp* sequence, a distance of 1156bp, which agrees well with the estimated size of 1.2kb for the cloned cDNA. Because of this, and the almost exact match of the pB43c1 sequences to the *usp* sequences, it was assumed that pB43c1 represents a partial *usp* cDNA, and that no further sequencing to confirm this was required, and that furthermore, the 43Y 3' transcriptional unit identified by reverse Northern is the gene *ultraspiracle*.

The *usp* gene product was cloned and sequenced independently by two different groups (Oro *et al.*, 1990; Henrich *et al.*, 1990). Both groups map the gene to cytological position 2C, with Henrich *et al.* locating it more specifically to 2C1-3. This agrees with the cytological location of the 43Y P[GAL4] insert which also maps to cytological position 2C. There is a single copy of the *usp* gene in the *Drosophila* genome, that produces a single intronless transcript. Both groups produced restriction maps of the *usp* genomic region, that are consistent with each other, and also consistent with the restriction map for the rescued plasmid pB43Y3'. Figure 7.5 is a comparison between the restriction maps for the *usp* genomic region and the rescued genomic DNA from pB43Y3'. The restriction maps place the P[GAL4] insertion 1.5kb downstream of the 3' end of the *usp* gene and in the opposite orientation.

The *usp* gene product (Usp) has homology to the retinoid X receptor class of proteins that are members of the nuclear hormone receptor superfamily. Usp interacts with the ecdysone receptor to produce a DNA binding heterodimer that mediates ecdysteroid-induced gene expression (Yao *et al.*, 1993). Developmental expression studies of *usp* revealed that it is expressed at all developmental stages including adult, with highest levels of transcripts and Usp protein in the late third larval instar (Henrich *et al.*, 1994). Immunohistochemistry on third instar larval tissues located Usp protein in the nuclei of many cell types including the central nervous system. Expression within the larval CNS occurs in all types of neurons, with no report of enhanced



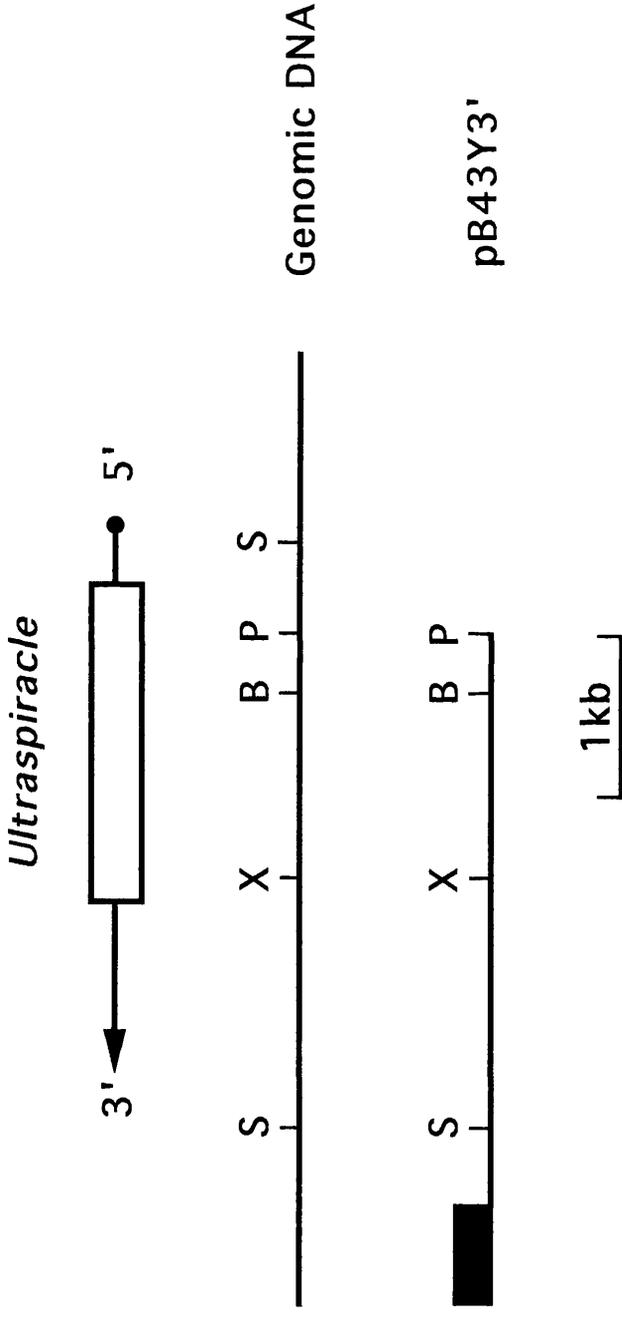


Figure 7.5. Comparison of restriction maps of the *usp* genomic region (adapted from Oro *et al.*, 1990) and the plasmid-rescued DNA of pB43Y3'. The position and orientation of the *usp* gene is indicated above the maps, the open box represents the translated region. The position of the extreme 3' end of the P[GAL4] insertion is indicated by the black box. Restriction enzymes: S = *SalI*; X = *XhoI*; B = *BamHI*; P = *PstI*.

expression in the larval mushroom bodies. Using an antibody (AB 11) I observed expression of Usp in the nuclei of all cells in the adult CNS, with no particular enhancement in the mushroom bodies (data not shown). It was therefore concluded that *ultraspiracle* gene expression does not reflect that of 43Y enhancer-trap expression, which is probably under the control of another gene's enhancer.

### **7.3 Analysis of the 43Y 5' gene**

#### **7.3.1 Isolation and sequence analysis of cDNA clones**

Plasmid rescue of 5' flanking genomic DNA, digested with restriction enzyme *KpnI*, yielded a plasmid containing 12.5 kb of genomic DNA known as pB43Y5'. Reverse Northern analysis of this plasmid identified a 4.0kb fragment containing a transcriptional unit. To facilitate identification of the gene involved, a cDNA library was screened using this 4.0kb fragment as a <sup>32</sup>P labelled probe. Approximately 100 000 recombinant phage plaques, from a *Drosophila* head cDNA library, were plated out for the primary screen. Ten positives were identified and purified down to single plaques. Phage DNA was prepared from each positive and digested with *EcoRI* and *Hind III* to release the cDNA insert. DNA fragments were separated by gel electrophoresis, Southern blotted onto nylon filters and probed with the 4.0kb fragment of pB43Y5' to confirm the positives as real (figure 7.6). From analysis of the restriction fragments the cDNAs can be split into six different classes (i.e. lanes 1 and 2, 3 and 4, 5 and 6, 7, 8, and 9 and 10, of the gel depicted in figure 7.6). Of these, all showed hybridisation to the pB43Y5' probe except lanes 3 and 4, and lane 7. Two cDNAs were selected for further detailed analysis, namely the cDNA of lane 8 (known as 7.1), and the cDNA of lane 9 (known as 8.2).

Initially simple restriction analysis was performed on cDNAs 7.1 and 8.2. Each cDNA was digested separately with *EcoRI*, and *HindIII*, and doubly digested with *EcoRI* and *HindIII*. The restriction fragments were separated by gel

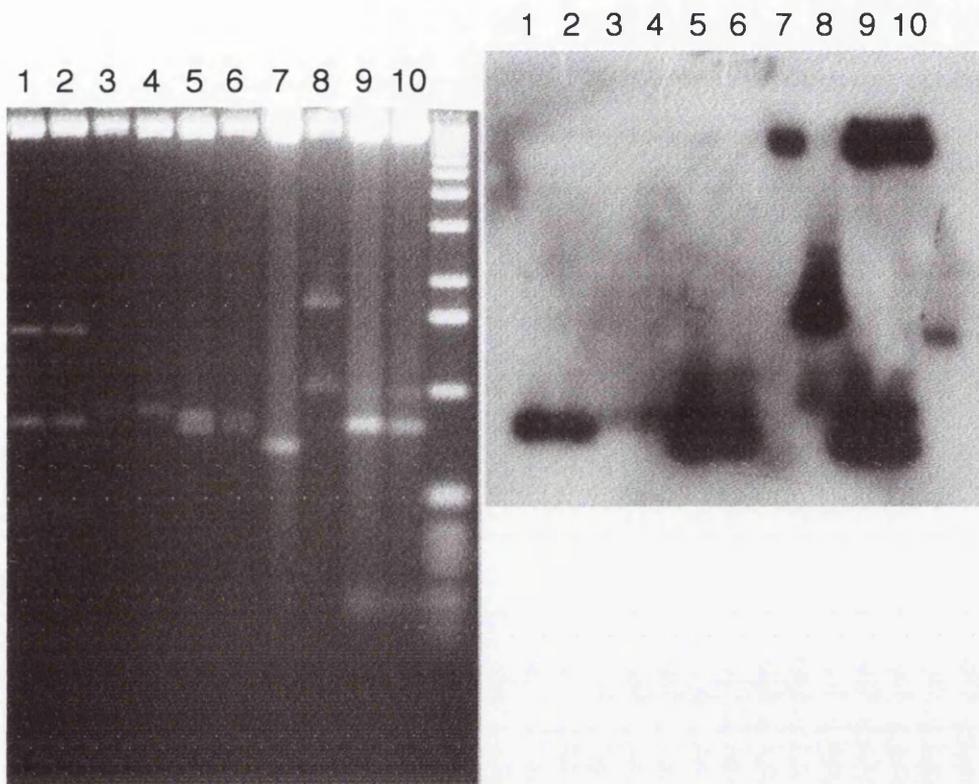


Figure 7.6. Isolated cDNAs, digested with *EcoRI* + *HindIII*. Gel blotted and probed with pB43Y5'. cDNAs corresponding to lanes 8 and 9 were selected for further analysis.

electrophoresis (figure 7.7), with the 7.1 cDNA yielding a 1.7kb *EcoRI* fragment and a 1.0kb *EcoRI/HindIII* fragment, whilst the 8.2 cDNA yielded a 0.3kb *EcoRI* fragment and a 0.9kb *EcoRI/HindIII* fragment. The 1.7kb *EcoRI* fragment and the 1.0kb *EcoRI/HindIII* fragment of cDNA 7.1 were subcloned into the vector pBluescript, creating new plasmids pB7.1A and pB7.1B respectively. Similarly, the 0.9kb *EcoRI/HindIII* fragment of cDNA 8.2 was cloned into pBluescript creating the new plasmid pB8.2. Using the primers T3 and T7, sequence data were obtained for the ends of the cDNA fragments cloned into plasmids pB7.1A, pB7.1B and pB8.2 (but no sequence was obtained from plasmid pB8.2 using primer T7). The sequence data acquired were used to search DNA databases using the BlastN algorithm. All sequences obtained (except the sequence from pB7.1B with primer T7) have homology to sequences from the *Drosophila*  $\alpha$ -actinin gene. The alignment between these sequences is shown in figure 7.8 (the few misalignments are again probably due to mistakes in the cDNA sequence obtained from the single sequencing runs). The sequence obtained from pB7.1B with primer T7 had no significant homology to any sequence in the databases, therefore just the plain sequence is given in figure 7.8. One would have expected this sequence to also have homology to  $\alpha$ -actinin sequence. According to two published maps of the  $\alpha$ -actinin genomic structure (Fyrberg *et al.*, 1990; Roulier *et al.*, 1992) the 3' UTR of the  $\alpha$ -actinin gene is approximately 500bp long. However, the  $\alpha$ -actinin sequence contained in Genbank document X51753 has only 57bp of 3' UTR. Assuming that pB7.1B contains a cDNA representing the extreme 3' end of the  $\alpha$ -actinin gene, then the 3' sequence of this cDNA (i.e. sequence obtained with the T7 primer) is part of the 3' UTR not included in the Genbank database. This would explain why the BlastN search found no homology for this sequence. The pB7.1B cDNA sequence starts from nucleotide 2323 of the  $\alpha$ -actinin sequence, the coding region of which extends to nucleotide 2754, and with a 3' UTR of 500bp gives a distance of 931bp. This agrees well with the size of the pB7.1B cDNA. I therefore have assumed that the pB7.1B cDNA represents the 3' end of the  $\alpha$ -actinin gene.

7.1A T3: 131 ACTACTCGATGNAANNNGN...GNGTCGGNTGGGAGNAGNNGNTTACATCTATCAACNGNA 80  
 ||||| ||||| || | ||||| ||||| || | ||||| ||||| || | ||||| ||||| || |  
 2195 ACTACACGATGGAACCGCTGCCCGTCGGCTGGGAGCAGCTGCTTACATCTATCAACCCGCA 2254

7.1A T3: 74 ACATCAACGAGGTGGAGAACCAGATTCTCACCCCGNGACTCCAAGGGNATTAGCCAGGAGN 20  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

7.1A T3: 2255 ACATCAACGAGGTGGAGAACCAGATTCTCACCCCGCGACTCCAAGGGCATTAGCCAGGAGC 2314

7.1A T3: 14 AGCTGAANGAATTC 1  
 ||||| ||||| |||||

7.1A T3: 2315 AGCTGAACGAATTC 2328

7.1A T7: 1 ACAAGAACGTC AACNTACAGAACTTNCATGTGTCTCATTCAAGGATGG 46  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

7.1A T7: 557 ACAAGAACGTC AACCGTACAGAACTTCCATCTGTCTCGTTCAAGGATGG 602

7.1B T3: 1 GAATTCGCCAGCAGCTTCAACCCTTCGACAAGAACCCGCCACCCGCCCTGTCCGCCCGAG 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

7.1B T3: 2323 GAATTCGCCAGCAGCTTCAACCCTTCGACAAGAACCCGCCACCCGCCCTGTCCGCCCGAG 2382

7.1B T3: 61 GAGTTC AAGNCGTGCATGGNCTCGT 85  
 ||||| ||||| ||||| ||||| |||||

7.1B T3: 2383 GAGTTC AAGTCTGCTTGGTCTCGT 2407

```

7.1B T7 Seq.: 1 TTTTATGATRRNNTTGTGTTTATTAACATAAATTCNAAAAATCCGTTTANCITTTT 60
                61 TTTAGCGTNTTTGGACNNCSTTTTTGTAAATAGATTTAAACCTTCGAT 109

8.2 T3:      1 GAATCCNCAGCAGCTTCAACCACCTTCGACAAGAACCACCGCCGACTGTCCGCCGAG 60
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
α-actinin: 2323 GAATCCGCAGCAGCTTCAACCACCTTCGACAAGAACCACCGCCGACTGTCCGCCGAG 2382

8.2 T3:      61 GAGTT...GT.GTGCTTNGTCTCGTTGGGCTACTCGATTGGCAAGGATCGCCARGGTGAT 116
                ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
α-actinin: 2383 GAGTTCAAAGTCTGCTGGTCTCGTTGGGCTACTCGATTGGCAAGGAAACGCCAGGGTGAT 2442

8.2 T3:      117 TTGGACTTNCANCNCATNCNC 137
                ||||| ||| ||| |||
α-actinin: 2443 TTGGACTTCCAGCGCATCCTC 2463

```

Figure 7.8. Alignment of sequences from ends of cDNA subclones pB7.1A, pB7.1B and pB8.2, using T3 and T7 primers, to  $\alpha$ -actinin sequences. The plain sequence from pB7.1B, using the T7 primer, is also given.  $\alpha$ -Actinin numbers correspond to those used in Genbank document X51753.

### 7.3.2 Restriction mapping of pB43Y5'

The *Drosophila*  $\alpha$ -actinin gene was first cloned and sequenced by Fyrberg *et al.*, 1990, and mapped to cytological location 2C, which again agrees with the position of the 43Y P[GAL4] insertion point. As part of their analysis of  $\alpha$ -actinin gene function they produced *EcoRI*, *HindIII*, and *BamHI* restriction maps for the  $\alpha$ -actinin genomic region. Therefore, to verify that  $\alpha$ -actinin is the 43Y 5' gene, and to determine its distance from, and orientation to, the P[GAL4] insertion point, restriction digests with enzymes *EcoRI*, *HindIII* and *BamHI*, and combinations thereof, were performed on the rescued plasmid pB43Y5'. Restriction fragments were separated by gel electrophoresis and the results are depicted in figure 7.9. Presumably, if the 43Y P[GAL4] insertion point is close to the  $\alpha$ -actinin locus, then pB43Y5' will have restriction fragments in common with those of the  $\alpha$ -actinin genomic region. This is indeed the case, examination of figure 7.9 reveals that pB43Y5' has the following *EcoRI* restriction fragments: 0.5kb, 1.2kb, 1.8kb, 2.1kb, and 4.0kb, all of which form a single contiguous stretch of the published genomic map. Similarly, the following pB43Y5' *HindIII* fragments are present: 0.5kb, 1.3kb, and 7.3kb, which again form a single contiguous stretch of the genomic map. And finally, the following pB43Y5' *BamHI* fragments are present: 2.3kb and 4.5kb, which form a single contiguous stretch of the genomic map. All the identified restriction fragments form a contiguous stretch towards the 3' end of the  $\alpha$ -actinin gene, and further analysis of the pB43Y5' restriction fragments, together with the information contained in the published map, and the known sites within the P[GAL4] element, allows the maps depicted in figure 7.10 to be constructed. The restriction maps place the 43Y P[GAL4] element approximately 6.5kb from the 3' end of the  $\alpha$ -actinin gene and in the same orientation. As a further check the right hand gel of figure 7.9 was Southern blotted onto nylon filters and probed with the cDNA from pB7.1B labelled with <sup>32</sup>P. This cDNA covers the  $\alpha$ -actinin gene from the *EcoRI* site to the end of the 3' UTR, and should hybridise to a 2.1kb *EcoRI* fragment, a 7.3kb *HindIII* fragment and a 4.5kb *BamHI* fragment. The

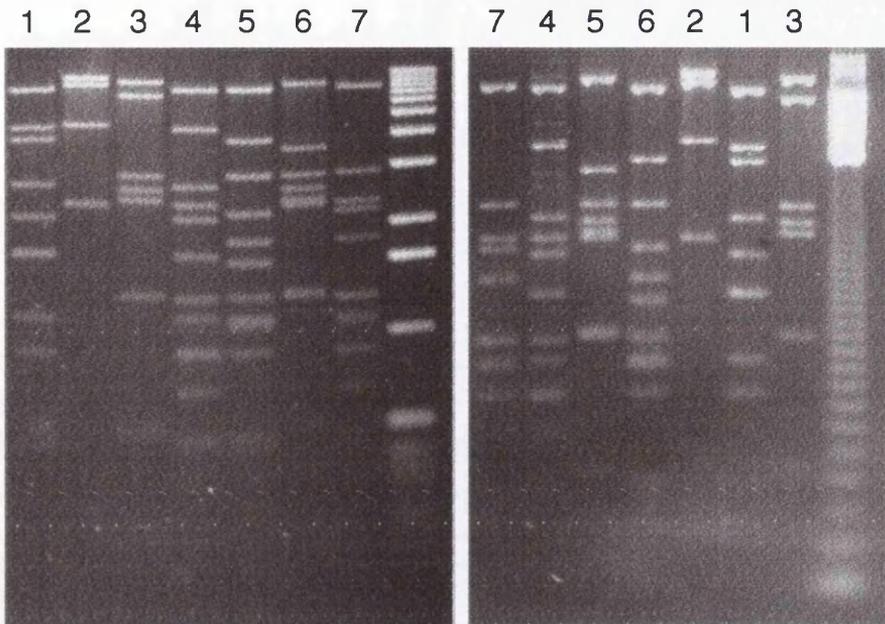


Figure 7.9. Restriction digests of pB43Y5'. Both gels show the same digests. Lane 1 = *EcoRI*, lane 2 = *BamHI*, lane 3 = *HindIII*, lane 4 = *BamHI* + *EcoRI*, lane 5 = *EcoRI* + *HindIII*, lane 6 = *BamHI* + *HindIII*, lane 7 = *EcoRI* + *BamHI* + *HindIII*. Fragments sized with BRL 1kb and 123bp ladders.

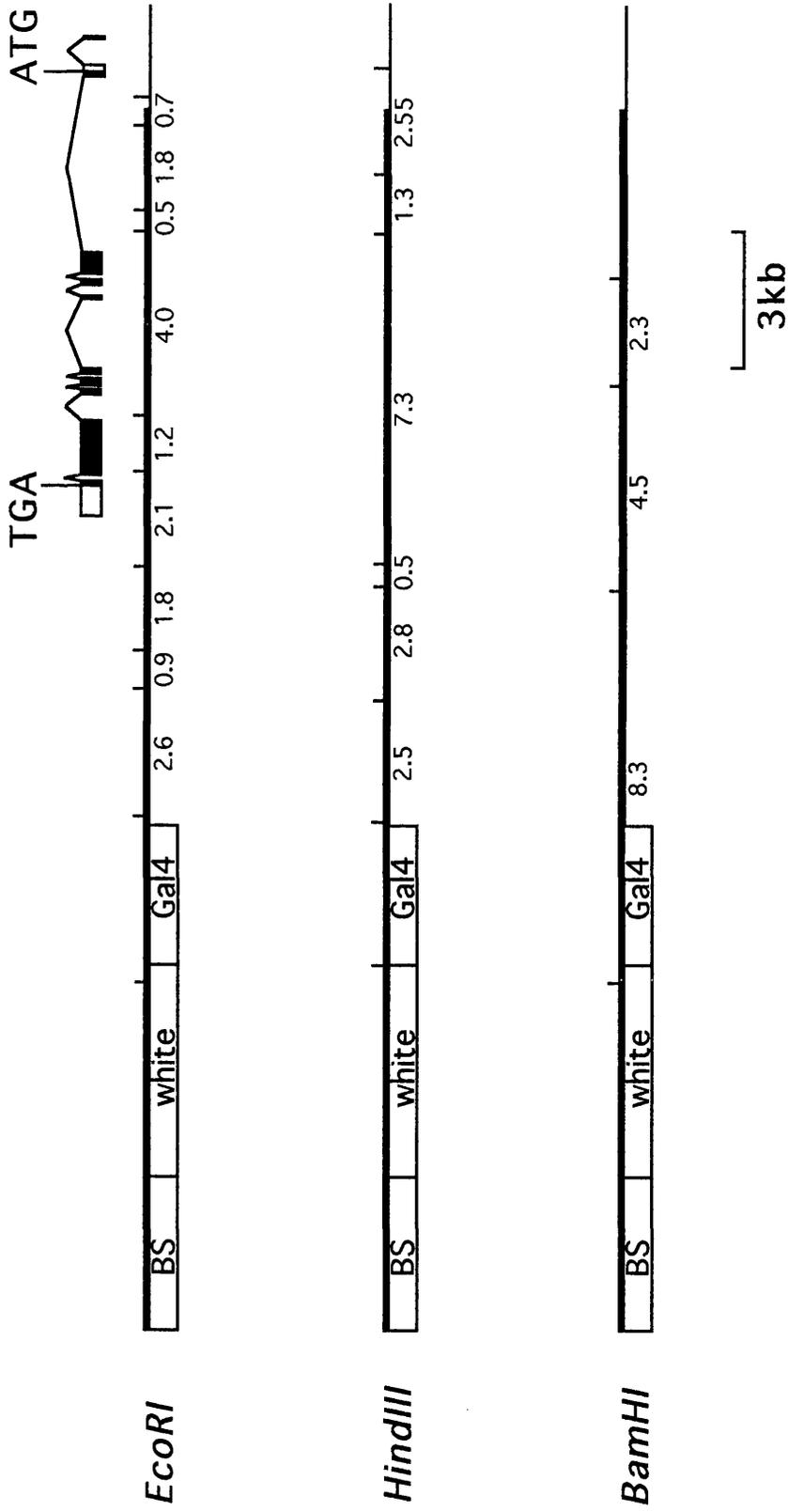


Figure 7.10. Restriction maps of the genomic region 5' to the 43Y P[GAL4] insertion. Size of fragments in kb are indicated. Position and orientation of the  $\alpha$ -actinin gene is indicated. Extent of the 43Y 5' rescued-plasmid is indicated by thick black line. (Based partly on data in Fyberg *et al*, 1990).

result is depicted in figure 7.11, and the expected fragments do hybridise to the probe, with fragments from multiple digests also hybridising in a predictable manner. The results outlined in this section indicate that the transcriptional unit identified in the rescued-plasmid pB43Y5', by reverse Northern, is the  $\alpha$ -actinin gene.

### 7.3.3 Expression of $\alpha$ -actinin

Does  $\alpha$ -actinin expression reflect that of the 43Y enhancer-trap expression? To answer this question a monoclonal antibody, MAC 276, raised against  $\alpha$ -actinin from the water bug *Lethocerus indicus* was acquired (gift from G. Butcher, Dept. Immunology, Babraham Institute; Lakey *et al.*, 1990). This same antibody was used by Roulier *et al.*, 1992, in their studies on *Drosophila*  $\alpha$ -actinin function, and specifically binds to *Drosophila*  $\alpha$ -actinin. Whole brains were dissected from adult *Drosophila* and immunohistochemistry performed with the  $\alpha$ -actinin antibody as outlined in Chapter 2. Images were acquired by confocal microscopy and results are shown in figure 7.12. Expression in the  $\alpha$ -lobe and  $\beta/\gamma$ -lobes of the mushroom bodies can clearly be seen, with expression in other areas of the brain being at a much lower level. Control experiments, in which the  $\alpha$ -actinin antibody was omitted (but including the labelled secondary antibody), showed no specific or elevated staining of the mushroom bodies. These results clearly suggest that  $\alpha$ -actinin has enhanced expression in the mushroom bodies, and that enhancers that control  $\alpha$ -actinin expression direct P[GAL4] expression in the fly line 43Y. If this is the case 43Y P[GAL4] expression should be observable in adult body structures that are known to express  $\alpha$ -actinin, such as the indirect flight muscles. Whole-body sections were acquired from fly line 43Y and stained for P[GAL4] directed  $\beta$ -galactosidase expression as described in chapter two. Results are shown in figure 7.13. Expression can clearly be seen in the thoracic indirect flight muscles, as well as in other structures such as gut muscles, the abdominal retractor of sternite muscles, and the wing hinge region. These other structures would also reasonably be expected to express  $\alpha$ -actinin. This clearly lends

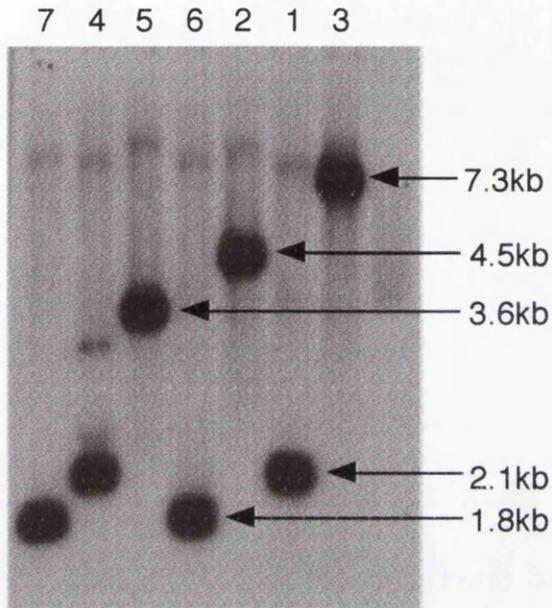


Figure 7.11. Right hand gel from figure 7.9 Southern blotted and probed with the cDNA from pB7.1B. Sizes of hybridising fragments are indicated. Lane numbers the same as for figure 7.8.

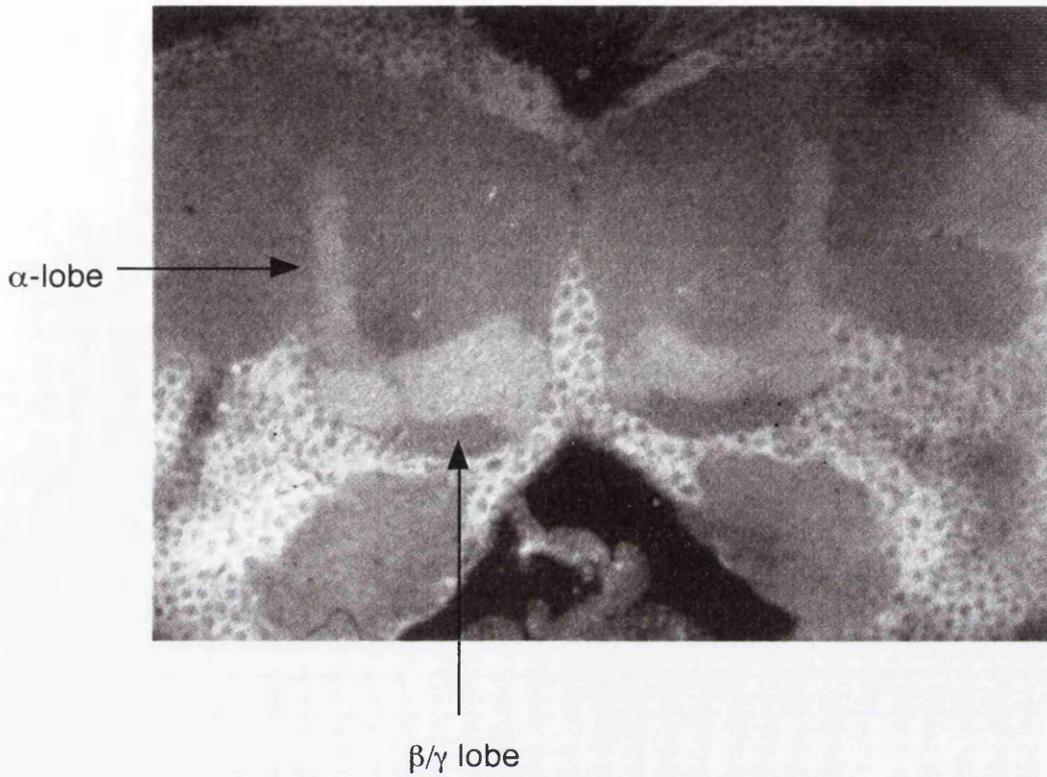


Figure 7.12. Confocal image of *Drosophila* adult whole brain, stained with a monoclonal antibody to  $\alpha$ -actinin. Staining in the  $\alpha$  and  $\beta/\gamma$  lobes of the mushroom bodies can clearly be seen.

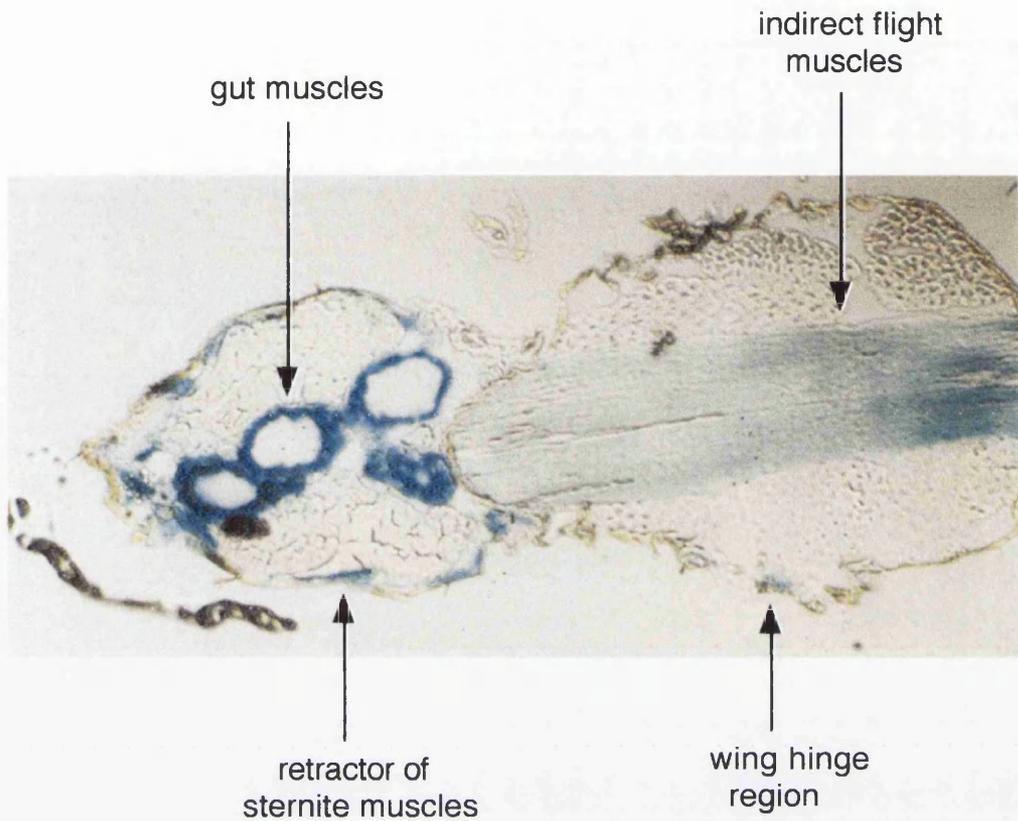


Figure 7.13. Horizontal section through the body of an adult fly from fly-line 43Y, stained for  $\beta$ -galactosidase expression. Staining can clearly be seen in the indirect flight muscles, gut muscles, abdominal retractor of sternite muscles, and the wing hinge region.

considerable weight to the thesis that enhancers, that control  $\alpha$ -actinin expression, direct 43Y P[GAL4] expression, and that  $\alpha$ -actinin has enhanced expression in the mushroom bodies.

## 7.4 Discussion

Analysis of the genomic region adjacent to the 43Y P[GAL4] insertion point, presented in this chapter, identified two genes that have previously been cloned and isolated in *Drosophila*, namely  $\alpha$ -actinin and *ultraspiracle*, and for the first time reveals the close proximity of these two genes. Furthermore, the evidence points clearly to  $\alpha$ -actinin expression being elevated in the mushroom bodies, and reflecting that of 43Y enhancer-trap expression.

### 7.4.1 *Drosophila* $\alpha$ -actinin

The *Drosophila*  $\alpha$ -actinin gene has a single genomic copy that produces three alternatively spliced mRNAs (Roulier *et al.*, 1992). A schematic diagram of the different mRNA isoforms is presented in figure 7.14. The principal difference between the particular mRNAs is the inclusion of the larval muscle, adult muscle, or non-muscle specific exon. The non-muscle isoform is expressed in early embryos, which lack developing myoblasts, and early pupae which have little muscle tissue. It is not known which isoforms are expressed in neurons in general or in the mushroom bodies in particular.

A mutation, *fliA*<sup>3</sup>, that specifically reduces expression of the adult muscle isoform, results in paralysis and atrophy of flight muscles, and weakening of leg muscles. However, the mutation, *fliA*<sup>4</sup>, that greatly reduces expression of all isoforms, has the same muscle defects but no obvious nonmuscle phenotype; embryo-genesis is normal as is adult morphology, suggesting that any embryogenic function of the nonmuscle isoform is redundant.

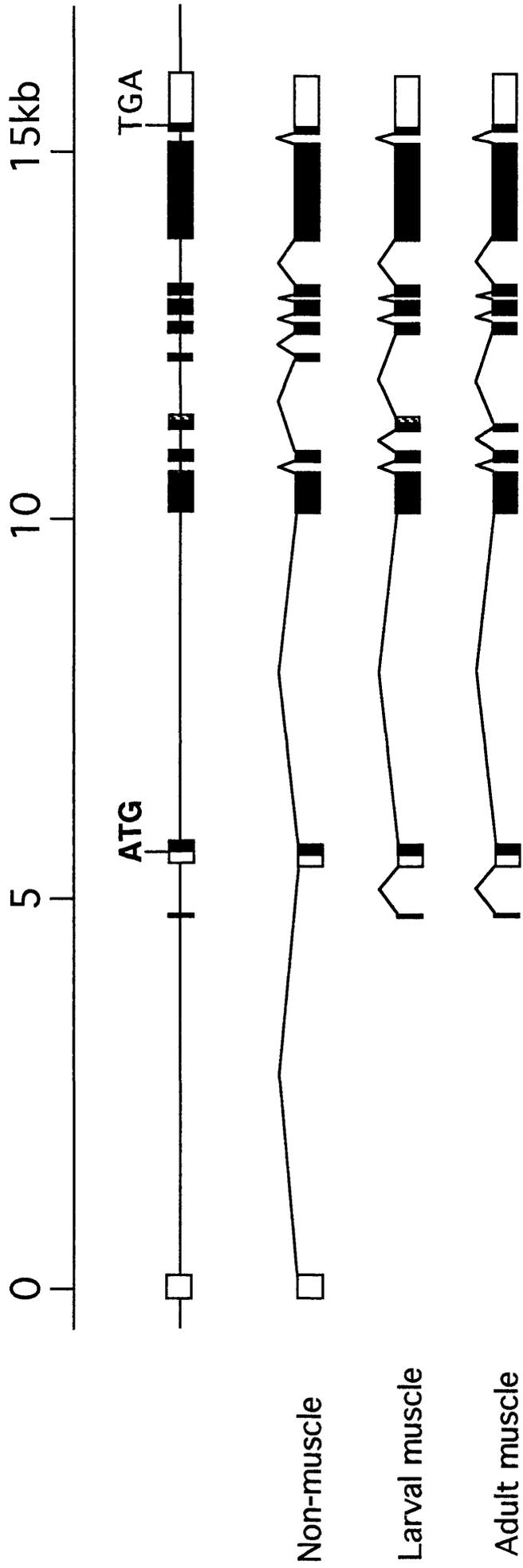


Figure 7.14. Schematic representation of alternatively spliced *Drosophila*  $\alpha$ -actinin mRNAs. The upper line illustrates the chromosomal arrangement of exons in the *Drosophila*  $\alpha$ -actinin gene. Solid and hatched blocks represent protein coding regions, while open blocks denote untranslated regions. (Adapted from Roulier *et al*, 1992).

All  $\alpha$ -actinins contain three distinct domains, an amino-terminal actin binding domain, a central repeat region, and a carboxy-terminal domain containing two EF-hand-like sequences. The *Drosophila* adult muscle isoform is a protein of 896 residues that has 68% overall identity to chicken  $\alpha$ -actinin. The actin binding domains of the two species are 86% identical, with the central repeats also being highly conserved (Fyrberg *et al.*, 1990).

*Drosophila*  $\alpha$ -actinin contains two copies of EF-hand-like domains, that are thought to bind  $\text{Ca}^{2+}$ . Figure 7.15 shows a comparison between the *Drosophila* EF-hand-like sequences and those from other species. Both *Drosophila* EF-hand-like domains have sufficient residues in accord with rules for calcium binding loops, outlined by Kretsinger, (1980), to suggest that they should bind  $\text{Ca}^{2+}$  *in vivo*. Although there is no *in vivo* or *in vitro* evidence for  $\text{Ca}^{2+}$  binding by *Drosophila*  $\alpha$ -actinin, the significance of this possibility is important to potential neuronal function, and will be discussed more fully later.

#### 7.4.2 $\alpha$ -Actinin neuronal function

$\alpha$ -Actinin is, in its dimeric form, a protein that cross-links actin filaments to form actin bundles, and is commonly associated with the Z-discs of muscle fibres (Blanchard *et al.*, 1989). It may therefore seem paradoxical that this protein also has elevated expression within the mushroom body structures of the *Drosophila* brain. Most research on  $\alpha$ -actinin has concentrated on its role in muscle function. However, actin and  $\alpha$ -actinin, as well as other actin binding proteins, are not restricted to muscle fibres but form a diverse array of structures in every cell. In particular, a network of actin microfilaments just under the surface of the cell membrane, forms an integral part of the cytoskeleton within neurons. It is this actin cytoskeleton that forms the key to understanding the potential function of  $\alpha$ -actinin in the mushroom bodies. This neuronal cytoskeleton is composed of three types of polymeric filaments, namely microtubules, neurofilaments, and actin micro-filaments. Microtubules and neuro-

```

* * * * *
EL LL LD D DG ID EL LL L          EL LL LD D DG ID EL LL L
Dictyostelium EFKACFSHFDKNDNKNLNRLEFSSCLKSIGDEL-----TEEQLNQVISKIDTDGNGTISFEEFIDYMVSS

Drosophila EFRSSEFNHFDKNRTGRLSPEEFKSCLVSLGYSIGKERQGDLDFQRILAVVDPNPNTGYVHFDAFLDFMTRE

Chicken NM  EFRASFNHFDRDHSGLGPEEFKACLISLGYDIGNDAQGEAEFARIMSIVDPNRMGVVTFQAFIDFMSRE

Chicken SM  EFRASFNHFDRKKTGMDCEDFRACLISMGYNM-----GEAEFARIMSIVDPNRMGVVTFQAFIDFMSRE

Chicken SkM DFRASFNHFDRRKNGLMDHDDFRACLISMGYDL-----GEAEFARIMSLVDPNGQGTVTFQSFIDFMTRE

```

Figure 7.15. Comparisons of EF-hand-like sequences of *Dictyostelium*, *Drosophila*, and chicken nonmuscle (NM), smooth muscle (SM), and skeletal muscle (SkM) isoforms of  $\alpha$ -actinin. The 16 positions where amino acid identity is requisite for an EF-hand structure are indicated above the sequences. Positions denoted by "E" are typically glutamic acid, "L" can be phenylalanine, leucine, isoleucine, valine, methionine, alanine, or tyrosine, "D" can be aspartic acid, asparagine, glutamic acid, glutamine, serine, or threonine, "G" is typically glycine, and "I" is isoleucine or valine. Asterisks denote positions where oxygen containing residues can coordinate calcium. *Drosophila* residues that match the required consensus are underlined. Functional EF-hands have correct residues at 12 or more of the 16 positions. (Adapted from Fyrberg *et al*, 1990).

filaments are oriented longitudinally in axons and dendrites, and are crosslinked to each other as well as to filaments of the same type by accessory proteins. Actin microfilaments form a meshwork underneath the membrane of the entire neuron, but are particularly enriched at the periphery of axons and in growth cones.

The neuronal cytoskeleton functions as a highway for molecular transport, and gives mechanical strength and shape to neuronal processes. Furthermore, the cytoskeleton is not merely static 'scaffolding' but is a dynamic structure that plays a central role in growth cone movement, neurite extension and in the generation and maintenance of axonal and dendritic arbours during development. Changes in neuronal architecture, in response to changing function, are also largely mediated by the cytoskeleton. By interacting with membrane proteins, to restrict their localisation to appropriate domains, the cytoskeleton plays a major role in establishing and maintaining the regional specialisation within neurons. It is clear that the cytoskeleton plays a critical and active role in neuronal function.

Actin microfilament fibres (F-actin) are polymers composed of 43 kD globular monomeric subunits (G-actin). Two F-actin strands twist around each other to form a double helical two-stranded rope. Numerous cytoplasmic factors are involved in the polymerisation process whereby G-actin units string together to form F-actin. Other factors form cross-links between actin fibres, bundle them together into parallel skeins and/or act as 'spacers', keeping parallel bundles apart by distances of about 200nm. In the axon the micro filaments are usually quite short, seldom more than about 0.5mm in length. Self-assembly of actin requires nucleoside triphosphates and proceeds until the monomer concentration is reduced to a critical concentration. At the critical concentration, assembly and disassembly are balanced. Above this concentration, there is net addition to the polymer, and below it, net removal. Actin microfilaments are polar structures, with one end (termed the plus end) growing faster during polymerisation than the other end (termed the minus end). Actin binds ATP which is hydrolysed to ADP upon polymerisation. ATP-actin is added primarily to the plus end, while ADP-actin readily dissociates when exposed from the minus end. Because

polymerisation and dissociation occur by different mechanisms, their critical concentrations can be different. At a monomer concentration between these two critical concentrations actin subunits are added at one end and lost at the other at the same rate. Thus, at steady state, the polymer treadmills, remaining approximately the same length as subunits move through it. This treadmilling is a potential source of mechanical energy whose ultimate source is the hydrolysis of ATP.

Actin in neurons exists in both monomeric and polymeric forms, the balance between the two being carefully controlled by the cell. The equilibrium between assembly and disassembly of the actin polymer can be shifted in response to a variety of environmental cues, such as interactions with growth factors or the extracellular matrix. Actin microfilaments are not longitudinally aligned along axons but form a complex meshwork beneath the surface membrane. The filaments are connected by numerous actin-binding proteins that bundle them together and cross-link them to form a gel. Other proteins control the length of microfilaments by binding to their ends and preventing further polymerisation, or by cleaving them into filaments of smaller length. Yet other proteins attach microfilaments to the surface membrane or to cytoskeletal proteins such as fodrin or vinculin.

The actin cytoskeleton plays a crucial role in growth cone motility and actin microfilaments comprise the majority of the cytoskeleton in growth cones (Smith, 1988). Actin filaments in the filopodia and lamellipodia of growth cones are in a highly dynamic state, and are predominantly oriented with their faster growing ends at the periphery and their slower growing ends towards the centre of the cell (Lewis and Bridgman, 1992). Guidance of the growth cone appears to depend mainly on actin microfilaments, while elongation is more fundamentally dependent on microtubules (Letourneau *et al.*, 1987).

A number of second messengers have a role in regulating growth cone motility, and in particular calcium appears to be a key regulatory agent. Intracellular calcium within neurons is increased by influx through voltage-gated and ligand-gated  $\text{Ca}^{2+}$  channels and by release from intracellular stores, and is lowered by membrane  $\text{Ca}^{2+}$ -

ATPases and Na<sup>+</sup>/Ca<sup>2+</sup> exchangers. Experiments on primary cultured neurons revealed that increasing intracellular calcium causes growth cone collapse and neurite retraction, presumably with a concomitant disassembly of the microfilament network. However, in neuronal tumour cell lines and cultured retinal neurons, treatment to induce an increase in Ca<sup>2+</sup> influx enhances neurite outgrowth and neurite extension. These observations may be reconciled by evidence that suggests the growth cone is highly sensitive to a narrow range of Ca<sup>2+</sup> levels (Lankford and Letourneau, 1991), and the proposal that the Ca<sup>2+</sup> optimum for different components in growth cones could differ (Kater *et al.*, 1988).  $\alpha$ -Actinin is potentially sensitive to calcium via its EF-hand motifs, and as well as cross-linking actin filaments links the actin cytoskeleton to integrins, receptor proteins which mediate growth cone recognition and attachment to molecules in the extracellular matrix (Otey *et al.*, 1990). Ca<sup>2+</sup>-sensitive  $\alpha$ -actinin in association with actin filaments has been identified in filopodia, and it has been suggested that some of the essential factors in Ca<sup>2+</sup>-dependent filopodia motility are the  $\alpha$ -actinin regulation of actin bundle formation, and the  $\alpha$ -actinin crosslinking of actin filaments to integrins (Sobue and Kanda, 1989).

#### **7.4.3 $\alpha$ -Actinin and NMDA receptors**

Another potentially important neuronal function for  $\alpha$ -actinin may be in the regulation of the N-methyl-D-aspartate (NMDA) class of ionotropic glutamate receptors. The activity of NMDA receptors is known to be dependent on the integrity of the actin cytoskeleton (Rosenmund and Westbrook, 1993). In particular, calcium-dependent rundown of NMDA channels, in cultured hippocampal neurons, is prevented when actin depolymerisation is blocked by phalloidin. This effect is specific to NMDA receptors as the Ca<sup>2+</sup>-dependent rundown of other ionotropic glutamate receptors is unaffected by phalloidin. It is suggested that a hypothetical calcium-binding regulatory protein links the actin cytoskeleton to NMDA receptors. Dissociation of the regulatory protein from NMDA receptors leads to their inactivation and run-down. A recent report

suggests that  $\alpha$ -actinin may be this hypothetical regulatory protein, and that it mediates the functionally important interaction between NMDA receptors and the post-synaptic actin cytoskeleton (Wyszynski *et al.*, 1997). NMDA receptors exist *in vivo* as heteromultimeric complexes consisting of the essential NR1 subunit coassembled with various subunits of the NR2 subfamily. A yeast two-hybrid screen identified  $\alpha$ -actinin as a protein that interacts directly, through its central rod domain, with the NR1 and NR2B subunits of the NMDA receptor. Furthermore,  $\alpha$ -actinin is coimmunoprecipitated with the NR1 and NR2B subunits, and is colocalised with NR1, by fluorescence immunohistochemistry, to dendritic spines in rat brain neurons. Interestingly,  $\text{Ca}^{2+}$ /calmodulin which binds to, and inhibits the activity of, the NMDA receptor, can antagonise  $\alpha$ -actinin binding to the NR1 subunit in a  $\text{Ca}^{2+}$ -dependent manner. This suggests that  $\text{Ca}^{2+}$ /calmodulin may displace  $\alpha$ -actinin from NR1 subunits in response to postsynaptic  $\text{Ca}^{2+}$  influx, perhaps through activated NMDA receptors themselves. This could lead to  $\text{Ca}^{2+}$ -dependent detachment of NMDA receptors from the actin cytoskeleton and may be a factor in the down regulation of NMDA receptors following  $\text{Ca}^{2+}$  influx (Tong *et al.*, 1995).  $\text{Ca}^{2+}$ /calmodulin does not bind to the NR2B subunit and thus NMDA receptors with distinct NR2 subunits may be differentially regulated by  $\text{Ca}^{2+}$ /calmodulin and the cytoskeleton.

NMDA receptors have been shown to be important in the establishment of associative long-term potentiation (LTP). Application of NMDA antagonists completely blocks the ability to induce LTP in synapses of the perforant path and the Schaffer collateral pathway of the mammalian hippocampus (Morris *et al.*, 1986). The NMDA receptor is voltage-regulated as the open channel is occluded at normal resting potential by a magnesium ion. Depolarisation drives the magnesium ion out of the channel, allowing monovalent cations and  $\text{Ca}^{2+}$  to pass. Thus, glutamate is most effective in opening the NMDA channel when the cell is depolarised. It is this voltage-dependent property of NMDA receptors that underlie its role in the formation of associative LTP. At resting membrane potentials, the ion channels linked to NMDA receptors are partially blocked by the normal extracellular concentration of  $\text{Mg}^{2+}$ .

However, when the membrane is depolarised, either during high-frequency tetanus or by simultaneous stimulation of neighbouring synapses, the affinity of the channels for  $Mg^{2+}$  decreases and the  $Mg^{2+}$  blockade is relieved. Under these circumstances, glutamate released from the presynaptic terminal is able to open NMDA receptor coupled channels, allowing  $Na^+$  and  $Ca^{2+}$  to flow into, and  $K^+$  to flow out of, the postsynaptic terminal (figure 7.16). The rise in  $Ca^{2+}$  provides a necessary trigger for subsequent events leading to LTP. Potential postsynaptic targets for  $Ca^{2+}$  include CaM kinase II and protein kinase C. Activity of both kinases is necessary to generate LTP (Malinow *et al*, 1989). It is probable that phosphorylation of membrane proteins by these kinases may be needed for the generation and maintenance of LTP.

Do NMDA receptors exist in *Drosophila*? Electrophysiology has revealed the existence of NMDA-gated cation channels in arthropods (Pfeiffer-Linn and Glantz, 1991). A cDNA encoding the R1 subunit of the NMDA receptor has been identified in *Drosophila* (Moriyoshi *et al*, 1991). The encoded protein (DNMDAR1) shows 46% overall identity with the rat NMDA R1 subunit, with a much higher identity of 79% for the putative membrane spanning domains that form the cation pore. However, expression of DNMDAR1 in *Xenopus* oocytes failed to generate a NMDA or glutamate dependent current. Seemingly other subunits are necessary for the assembly of functional NMDA-operated ion channels. Expression of DNMDAR1 is low in late embryos but reaches high levels in late pupae when the adult CNS is formed (Ultsch *et al*, 1993). DNMDAR1 expression continues in adulthood, and transcript levels are highest in RNA prepared from fly heads, consistent with its neuronal function and potential role in LTP.

## 7.5 Conclusion

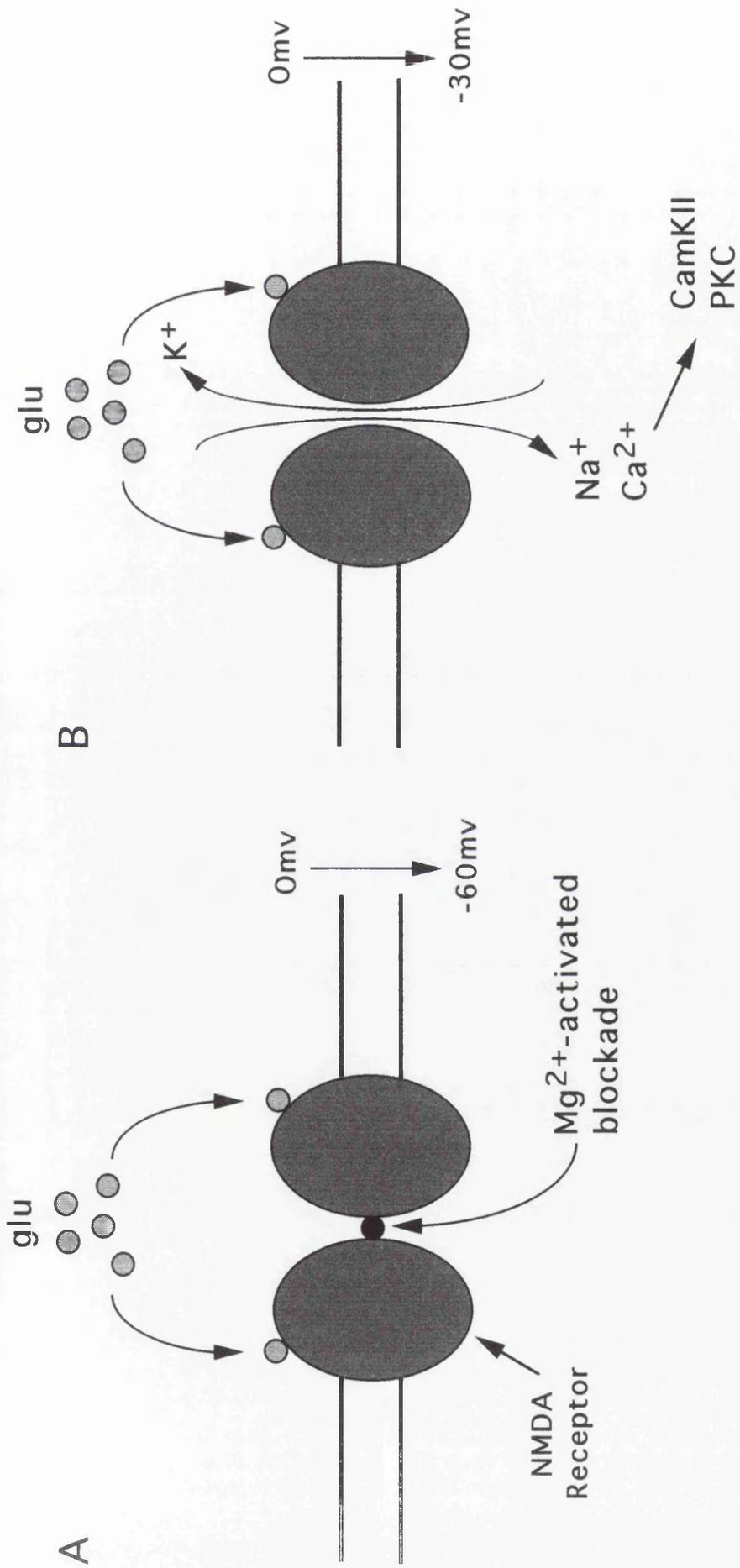


Figure 7.16. Mechanism of action of the NMDA receptor. (A) At normal resting potential the channel is blocked by a Mg<sup>2+</sup> ion. (B) Depolarisation releases the Mg<sup>2+</sup> blockade allowing glutamate to open the channel. The influx of Ca<sup>2+</sup> leads to the activation of CaM kinase II and protein kinase C.

It is clear from the previous discussion that there are a number of possible explanations for elevated expression of  $\alpha$ -actinin in the mushroom bodies of *Drosophila*. As outlined in chapter one there is good evidence that the mushroom bodies play a key role in *Drosophila* learning and memory. If this is the case, one can make the reasonable assumption that the mushroom bodies are dynamic structures, with a high degree of neuronal plasticity. Thus one can imagine that the intrinsic neurons of the mushroom bodies are constantly undergoing neurite outgrowth or retraction as new synapses are formed or lost in response to different environmental conditions or experiences. Interestingly, it has recently been shown that the growth cone motility of neurons cultured *in vitro*, from the *Drosophila* memory mutants *dunce* and *rutabaga*, is greatly reduced as compared to wild-type (Kim and Wu, 1996). It would appear that normal growth cone activity is important to the proper functioning of learning and memory in *Drosophila*. The role of  $\alpha$ -actinin in regulating growth cone activity, via its role in crosslinking actin microfilaments and/or by linking the actin microfilament network to integrins, seems particularly pertinent here. It seems reasonable to postulate that elevated  $\alpha$ -actinin expression in the mushroom bodies partially reflects the importance of  $\alpha$ -actinin in growth cone activity. However, more generally, the dynamic activity of mushroom body neurons would indicate a high degree of plasticity on the part of the actin microfilament network.  $\alpha$ -Actinin could obviously play a key role in regulating this plasticity, and its potential sensitivity to calcium places this regulation under the control of a key second messenger. On this basis the reduced level of  $\alpha$ -actinin expression in other neuronal structures suggests that they lack the dynamic activity of the mushroom bodies. Thus their intrinsic neurons are, compared to mushroom body intrinsic neurons, more or less 'hard-wired', and therefore the microfilament network does not require the degree of plasticity that  $\alpha$ -actinin confers.

The importance of LTP in the formation of memories has long been recognised, and the role of NMDA receptors in the establishment of LTP is well documented. The fact that  $\alpha$ -actinin may regulate NMDA receptor activity suggests another potential role for elevated  $\alpha$ -actinin expression in the mushroom bodies of *Drosophila*. Although it

appears that *Drosophila* have NMDA receptors it is not yet known whether they are expressed at a higher level in the mushroom bodies, as compared with other brain structures, or whether they interact directly with *Drosophila*  $\alpha$ -actinin. However, the possibility arises that  $\alpha$ -actinin plays a role in *Drosophila* learning and memory via its interaction with NMDA receptors in the mushroom bodies. This potentially important observation is worthy of further investigation.

## **Conclusions**

The aims of the research reported in this thesis were to identify genes with preferential expression in the mushroom bodies of *Drosophila melanogaster*. Identification of such genes is a first step in furthering our understanding of mushroom body physiology and biochemistry, and elucidating their role in fly learning and memory. Two fly lines, known as 201Y and 43Y, with enhancer-trap expression in the mushroom bodies were chosen for detailed study. Two genes flanking each of the two enhancer-traps were identified and analysed. Of these four genes, two had previously been identified in *Drosophila*, namely *ultraspiracle* and  $\alpha$ -actinin, one is a newly identified member of the FKBP12 family of proteins, and one, A421, is a new gene of unknown function.

From the reverse Northern and Northern analysis, gene A421 has a transcript that clearly has head elevated expression, and the *in situ* hybridisation data suggests this head expression is within neurons of the *Drosophila* brain. However, it would appear that this neuronal expression is not restricted to the mushroom bodies, and therefore does not reflect expression of the 201Y P[GAL4] enhancer-trap. Furthermore, the reverse Northern analysis of pB43Y3' indicated that *ultraspiracle* has head elevated expression, but immunohistochemistry reveals it to be ubiquitously expressed throughout the fly brain. It would appear that head elevated gene expression is not a good indicator of potential mushroom body expression, and that in retrospect the emphasis placed on this criterion in Chapter 3 was misplaced.

Although the A421 gene is not preferentially expressed in the mushroom bodies, analysis of the predicted protein sequence suggests it is a gene of some potential interest. The A421 protein's glutamine-rich region is suggestive of a transcriptional activation function. Close analysis of the A421 protein sequence revealed a novel protein motif, similar in structure to a C4 zinc-finger. Further work discovered this motif to exist in a large number of proteins of diverse functionality, and from a diverse array of species. No doubt many more proteins containing this motif will be uncovered in the future. Once the exact function of the motif is determined

experimentally, it may become, like other protein motifs, an important indicator of protein function.

The *Drosophila* FKBP12 homologue is very similar to FKBP12 proteins from other species indicating this a highly conserved family of proteins. Northern analysis indicated it to be expressed at all stages of *Drosophila* development. However, it was not possible to determine, using the acquired antibody, whether FKBP12 is preferentially expressed in the mushroom bodies, and further work using *in situ* hybridisation with DIG labelled cDNA probes, or other antibodies will be required. The possibility remains that FKBP12 expression reflects that of the 201Y enhancer-trap. However, the FKBP12 gene is located downstream of the enhancer-trap and in the opposite 5' to 3' orientation. Therefore, any upstream enhancers controlling FKBP12 expression will be a considerable distance from the *Gal4* gene of the enhancer-trap. This does of course not preclude the possibility that FKBP12 enhancers control 201Y enhancer-trap expression, and indeed the FKBP12 gene may be under the control of downstream enhancers.

The most important discovery of the research presented in this thesis is the elevated expression of  $\alpha$ -actinin in the mushroom bodies of *Drosophila*. Furthermore, I would suggest that this discovery vindicates the whole enhancer-trap approach of this thesis. Our current knowledge of mushroom body physiology and biochemistry would probably not have suggested  $\alpha$ -actinin as a good candidate for a gene with preferential mushroom body expression. Study of other fly lines with enhancer-trap expression in the mushroom bodies may uncover other similar genes that *a priori* analysis would have dismissed as 'unpromising'. As outlined in Chapter 7 there are, I think, two potential reasons why  $\alpha$ -actinin may be preferentially expressed in the mushroom bodies. Firstly, it could mediate, via the actin cytoskeleton, neuronal plasticity and growth cone activity. Or alternatively, it may regulate the activity of NMDA receptors. Both of these hypotheses indicate potentially fruitful avenues of research for furthering our understanding of mushroom body function.

## **Bibliography**

Aceves-Pina, E.O., and Quinn, W.G. (1979). Learning in normal and mutant *Drosophila melanogaster*. *Science* **206**, 93-96.

Alliegro, M.C., and Alliegro, M.A. (1996). A nuclear protein regulated during the transition from active to quiescent phenotype in cultured endothelial cells. *Dev. Biol.* **174**, 288-297.

Aperia, A., Ibarra, F., Svennson, L.-B., Klee, C., and Greengard, P. (1992). Calcineurin mediates  $\alpha$ -adrenergic stimulation of Na<sup>+</sup>, K<sup>+</sup>-ATPase activity in renal tubule cells. *Proc. Natl. Acad. Sci. USA* **89**, 7394-7397.

Armstrong, J.D. (1995). PhD Thesis, University of Glasgow.

Asztalos, Z., von Wegerer, J., Wustmann, G., Dombradi, V., Gausz, J., Spatz, H.-C., and Friedrich, P. (1993). Protein phosphatase-1-deficient mutant *Drosophila* is affected in habituation and associative learning. *J. Neurosci.* **13**, 924-930.

Balling, A., Technau, G.M., and Heisenberg, M. (1987). Are the structural changes in adult *Drosophila* mushroom bodies memory traces? Studies on biochemical learning mutants. *J. Neurogenetics* **4**, 65-73.

Ballinger, D.G., and Benzer, S. (1989). Targeted gene mutations in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **86**, 9402-9406.

Baltz, R., Evrard, J.-L., Domon, C., and Steinmetz, A. (1992). A LIM motif is present in a pollen-specific protein. *Plant Cell* **4**, 1465-1466.

Beddow, A.L., Richards, S.A., Orem, N.R., and Macara, I.G. (1995). The Ran/TC4 GTPase-binding domain: identification by expression cloning and characterization of a conserved sequence motif. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3328-3332.

Bellen, H.J., Develyn, D., Harvey, M., and Elledge, S.J. (1992). Isolation of temperature-sensitive diptheria toxins in yeast and their effects in *Drosophila* cells. *Development* **114**, 787-796.

Berg, J.M., and Shi, Y. (1996). The galvanization of biology: a growing appreciation for the roles of zinc. *Science* **271**, 1081-1085.

Berridge, M. (1993). Inositol triphosphate and calcium signalling. *Nature* **361**, 315-325.

Bierer, B.E., Somers, P.K., Wandless, T.J., Burakoff, S.J., and Schreiber, S.L. (1990). Probing immunosuppressant action with a nonnatural immunophilin ligand. *Science* **250**, 556-559.

Birnboim, H.C., and Doly, J. (1979). Rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* **7**, 1513-1523.

Blanchard, A., Ohanian, V., and Critchley, D. (1989). The structure and function of  $\alpha$ -actinin. *J. Muscle Res. Cell Motil.* **10**, 280-289.

Blumenthal, D.K., Takio, K., Hansen, R.S., and Krebs, E.G., (1986). Dephosphorylation of cAMP-dependent protein kinase regulatory subunit (type II) by calmodulin-dependent protein phosphatase. *J. Biol. Chem.* **261**, 8140-8145.

Borden, K.L.B., and Freemont, P.S. (1996). The RING finger domain - a recent example of a sequence-structure family. *Curr. Opin. Struct. Biol.* **6**, 395-401.

Boulikas, T. (1994). Putative Nuclear Localization Signals (NLS) in Protein Transcription Factors. *J. Cell. Biochem.* **55**, 32-58.

Bourgouin, C., Lundgren, S.E., and Thomas, J.B. (1992). *apterous* is a *Drosophila* LIM domain gene required for the development of a subset of embryonic muscles. *Neuron* **9**, 549-561.

Boynton, S., and Tully, T. (1992). *latheo*, a new gene involved in associative learning and memory in *Drosophila melanogaster*, identified from P element mutagenesis. *Genetics* **131**, 655-672.

Brand, A.H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fate and generating dominant phenotypes. *Development* **118**, 401-415.

Brilliantes, A.-M.B., Ondrias, K., Scott, A., Kobrinsky, E., Ondriasova, E., Landers, M., Ehrlich, B.E., and Marks, A.R. (1994). Stabilization of calcium release channel (ryanodine receptor) function by FK506-binding protein. *Cell* **77**, 513-523.

Brown, L., Chen, M.X., and Cohen, P.T.W. (1994). Identification of a cDNA encoding a *Drosophila* calcium/calmodulin regulated protein phosphatase, which has its most abundant expression in the early embryo. *FEBS Lett.* **339**, 124-128.

Bullock, W.O., Fernandez, J.M., and Short, J.M. (1987). XL1-Blue: A high efficiency plasmid transforming *recA Escherichia coli* strain with  $\beta$ -Galactosidase selection. *Biotechniques* **5**, 376-379.

Busch, S.J., and Sassone-Corsi, P. (1990). Dimers, leucine zippers and DNA-binding domains. *TIG* **6**, 36-40.

Byers, D., Davis, R.L., and Kiger, J.A., Jr. (1981). Defect in cyclic AMP phosphodiesterase due to the *dunce* mutation of learning in *Drosophila melanogaster*. *Nature* **289**, 79-81.

Cameron, A.M., Steiner, J.P., Sabatini, D.M., Kaplin, A.I., Walensky, L.D., and Snyder, S.H. (1995). Immunophilin FK506 binding protein associated with inositol 1,4,5-triphosphate receptor modulates calcium flux. *Proc. Natl. Acad. Sci. USA* **92**, 1784-1788.

Cavener, D.R. (1987). Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucl. Acids Res.* **15**, 1353-1361.

Chang, K.W., Laconi, S., Mangold, K.A., Hubchak, S., and Scarpelli, D.G. (1995). Multiple genetic alterations in hamster pancreatic ductal adenocarcinomas. *Cancer Res.* **55**, 2560-2568.

Chen, C.-n., Denome, S., and Davis, R.L. (1986). Molecular analysis of cDNA clones and the corresponding genomic coding region of the *Drosophila dunce*<sup>+</sup> gene, the structural gene for cAMP phosphodiesterase. *Proc. Natl. Acad. Sci. USA* **83**, 9313-9317.

Chen, M.X., Chen, Y.H., and Cohen, P.T.W. (1992). Polymerase chain-reactions using *Saccharomyces*, *Drosophila* and human DNA predict a large family of protein serine threonine phosphatases. *FEBS Lett.* **306**, 54-58.

Choi, K.-W., Smith, R.F., Buratowski, R.M., and Quinn, W.G. (1991). Deficient protein kinase C activity in *turnip*, a *Drosophila* learning mutant. *J. Biol. Chem.* **266**, 15999-16006.

Chung, K.W., Lee, Y.M., Park, T.K., Kim, S.J., and Lee, C.C. (1996). Cooperative Transcriptional Activation by Two Glutamine-rich Regions of *twist* Product in *Drosophila melanogaster*. *Mol. Cells* **6**, 197-202.

Coghlan, V.M., Perrino, B.A., Howard, M., Langeberg, L.K., Hicks, J.B., Gallatin, W.M., and Scott, J.D. (1995). Association of protein kinase A and protein phosphatase 2B with a common anchoring protein. *Science* **267**, 108-111.

Cohen, P. (1992). Signal integration at the level of protein kinases, protein phosphatases and their substrates. *TIBS* **17**, 408-413.

Cooley, L., Kelley, R., and Spradling, A. (1988). Insertional mutagenesis of the *Drosophila* genome with single P elements. *Science* **239**, 1121-1128.

Corfas, G., and Dudai, Y. (1989). Habituation and dishabituation of a cleaning reflex in normal and mutant *Drosophila*. *J. Neurosci.* **9**, 56-62.

Courey, A.J., and Tjian, R. (1988). Analysis of Sp1 In Vivo Reveals Multiple Transcriptional Domains, Including a Novel Glutamine-Rich Activation Motif. *Cell* **55**, 887-898.

Courey, A.J., Holtzman, D.A., Jackson, S.P., and Tjian, R. (1989). Synergistic Activation by the Glutamine-Rich Domains of Human Transcription Factor Sp1. *Cell* **59**, 827-836.

Cowan, T.M., and Siegel, R.W. (1986). *Drosophila* mutations that alter ionic conduction disrupt acquisition and retention of a conditioned odor avoidance response. *J. Neurogenet.* **3**, 187-201.

Crabtree, G.R. (1989). Contingent genetic regulatory events in lymphocyte-T activation. *Science* **243**, 355-361.

Crozat, A., Åman, P., Mandahl, N., and Ron, D. (1993). Fusion of CHOP to a novel RNA-binding protein in human myxoid lipo-sarcoma. *Nature* **363**, 640-644.

Dalby, D., Pereira, A.J., and Goldstein, L.S.B. (1995). An inverse PCR screen for the detection of P element insertions in cloned intervals in *Drosophila melanogaster*. *Genetics* **139**, 757-766.

Dash, P.K., Karl, K.A., Colicos, M.A., Prywes, R., and Kandel, E.R. (1991). cAMP response element-binding protein is activated by Ca<sup>2+</sup>/Calmodulin-dependent as well as cAMP dependent protein kinase. *Proc. Natl. Acad. Sci. USA* **88**, 5061.

Dauwalder, B., and Davis, R.L. (1995). Conditional rescue of the *dunce* learning/memory and female fertility defects with *Drosophila* or rat transgenes. *J. Neurosci.* **15**, 3490-3499.

Davis, R.L. (1993). Mushroom bodies and *Drosophila* learning. *Neuron* **11**, 1-14.

Davis, R.L., and Kiger, J.A., Jr. (1981). *dunce* mutants of *Drosophila melanogaster*-mutants defective in the cyclic AMP phosphodiesterase enzyme system. *J. Cell Biol.* **90**, 101-107.

Davis, R.L., and Davidson, N. (1986). The memory gene *dunce*<sup>+</sup> encodes a remarkable set of RNAs with internal heterogeneity. *Mol. Cell Biol.* **6**, 1464-1470.

Dawson, T.M., Steiner, J.P., Dawson, V.L., Dinerman, J.L., Uhl, G.R., and Snyder, S.H. (1993). Immunosuppressant FK506 enhances phosphorylation of nitric oxide synthase and protects against glutamate neurotoxicity. *Proc. Natl. Acad. Sci. USA* **90**, 9808-9812.

Dawson, T.M., Steiner, J.P., Lyons, W.E., Fotuhi, M., Blue, M., and Snyder, S.H. (1994). The immunophilins, FK506 binding protein and cyclophilin, are discretely localized in the brain: relationship to calcineurin. *Neurosci.* **62**, 569-580.

de Belle, J.S., and Heisenberg, M. (1994). Associative odor learning in *Drosophila* abolished by chemical ablation of mushroom bodies. *Science* **263**, 692-695.

Delaney, S.J., Hayward, D.C., Barleben, F., Fischbach, K.-F., and Gabor Miklos, G.L. (1991). Molecular cloning and analysis of small optic lobes, a structural brain gene of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, **88**, 7214-7218.

Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P., Rouleau, G., Aurias, A., and Thomas, G. (1992). Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359**, 162-165.

- Deleage, G., and Roux, B. (1987) An algorithm for protein secondary structure prediction based on class prediction. *Prot. Engng.* **1**, 289-294
- DiLella, A.G., and Craig, R.J. (1991). Exon organization of the human FKBP-12 gene: Correlation with structural and functional protein domains. *Biochemistry* **30**, 8512-8517.
- Drain, P., Folkers, E., and Quinn, W.G. (1991). cAMP-dependent protein kinase and the disruption of learning in transgenic flies. *Neuron* **6**, 71-82.
- Dudai, Y. (1983). Mutations affect storage and use of memory differentially in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **80**, 5445-5448.
- Dudai, Y., Sher, B., Segal, D., and Yovell, Y. (1985) Defective responsiveness of adenylate cyclase to forskolin in the *Drosophila* memory mutant *rutabaga*. *J. Neurogenet.* **2**, 365-380.
- Dudai, Y., Jan, Y.-N., Byers, D., Quinn, W., and Benzer, S. (1976). *dunce*, a mutant of *Drosophila* deficient in learning. *Proc. Natl. Acad. Sci. USA* **73**, 1684-1688.
- Dudai, Y., Buxbaum, J., Corfas, G., Orgad, S., Segal, D., Sher, B., Uzzan, A., and Zvi, S. (1986). Defective cAMP metabolism and defective memory in *Drosophila*. *Acta Biochim. Biophys. Hung.* **21**, 177-192.
- Duerr, J.S., and Quinn, W.G. (1982) Three *Drosophila* mutants that block associative learning also affect habituation and sensitization. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3646-3650.
- Dujardin, F. (1850). Memoire sur le system nerveux des insectes. *Ann. Sci. Nat. Zool.* **14**, 195-206.
- Dumont, F.J., Melino, M.R., Staruch, M.J., Koprack, S.L., Fisher, P.A., and Sigal, N.H. (1990). The immunosuppressive macrolides FK-506 and rapamycin act as reciprocal antagonists in murine T-cells. *J. Immunol.* **144**, 1418-1424.
- Dura, J.-M., Preat, T., and Tully, T. (1993). Identification of *linotte*, a new gene affecting learning and memory in *Drosophila melanogaster*. *J. Neurogenet.* **9**, 1-14.

Dura, J.-M., Taillebourg, E., and Preat, T. (1995). The *Drosophila* learning and memory gene *linotte* encodes a putative receptor tyrosine kinase homologous to the human RYK gene product. *FEBS Letts.* **370**, 250-254.

Engels, W.R., (1989) P elements in *Drosophila melanogaster*. In: *Mobile DNA* (American Society for Microbiology) pp. 437-484.

Engels, W.R., Benz, W.K., Preston, C.R., Graham, P.L., Phillis, R.W., and Robertson, H.M. (1987). Somatic effects of P element activity in *Drosophila melanogaster*: pupal lethality. *Genetics* **117**, 745-757.

Fakharzadeh, S.S., Trusko, S.P., and George, D.L. (1991). Tumorigenic potential associated with enhanced expression of a gene that is amplified in a mouse tumor cell line. *EMBO J.* **10**, 1565-1569.

Feany, M.B., and Quinn, W.G. (1995). A neuropeptide gene defined by the *Drosophila* memory mutant *amnesiac*. *Science* **268**, 869-873.

Ferreira, A., Kincaid, R., and Kosik, K.S. (1993). Calcineurin is associated with the cytoskeleton of cultured neurons and has a role in the acquisition of polarity. *Mol. Biol. Cell* **4**, 1225-1238

Ferreri, K., Gill, G., and Montminy, M. (1994). The cAMP-regulated transcription factor CREB interacts with a component of the TFIID complex. *Proc. Natl. Acad. Sci. USA* **91**, 1210-1213.

Fischer, G., and Schmid, F.X. (1990). The mechanism of protein folding - implications of *in vitro* refolding models for *de novo* protein folding and translocation in the cell. *Biochemistry* **29**, 2205-2212.

Fischer, J.A., Giniger, E., Maniatis, T., and Ptashne, M. (1988). GAL4 activates transcription in *Drosophila*. *Nature* **333**, 853-856.

Folkers, E., and Spatz, H.C. (1981). Visual learning behaviour in *Drosophila melanogaster*. *J. Insect Physiol.* **27**, 615-622.

Folkers, E., and Spatz, H.C. (1984). Visual learning performance of *Drosophila melanogaster* is altered by neuropharmaca affecting phosphodiesterase activity and acetylcholine transmission. *J. Insect Physiol.* **30**, 957-965.

Folkers, E., Drain, P., and Quinn, W.G. (1993). *radish*, a *Drosophila* mutant deficient in consolidated memory. Proc. Natl. Acad. Sci. USA **90**, 8123-8127.

Freemont, P.S., Hanson, I.M., and Trowsdale, J. (1991). A novel cysteine rich sequence motif. Cell **64**, 483-484.

Furuichi, T., Furutumi, D., Hakamata, Y., Nakai, J., Takeshima, H., and Mikoshiba, K. (1994). Multiple types of ryanodine receptor  $Ca^{2+}$  release channels are differentially expressed in rabbit brain. J. Neurosci. **14**, 4794-4805.

Fyrberg, E., Kelly, M., Ball, E., Fyrberg, C., and Reedy, M.C. (1990). Molecular Genetics of *Drosophila* Alpha-Actinin: Mutant alleles disrupt Z disc integrity and muscle insertions. J. Cell Biol. **110**, 1999-2011.

Galat, A. (1993). Peptidylproline *cis-trans*-isomerases: immunophilins. Eur. J. Biochem. **216**, 689-707.

Geourjon, C., and Deleage, G. (1994) SOPM - A self-optimized method for protein secondary structure prediction. Prot. Engng., **7**, 157-164.

Gerber, H.-P., Katja, S., Georgiev, O., Höfferer, M., Hug, M., Rusconi, S., and Schaffner, W. (1994). Transcriptional Activation Modulated by Homopolymeric Glutamine and Proline Stretches. Science **263**, 808- 811.

Gibrat, J.F., Garnier, J., and Robson, B. (1987) Further developments of protein secondary structure prediction using information theory - new parameters and consideration of residue pairs. J.Mol.Biol. **198**, 425-443

Gill, G., Pascal, E., Tseng, Z.H., and Tjian, R., (1994). A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAF<sub>II</sub>110 component of the *Drosophila* TFIID complex and mediates transcriptional activation. Proc. Natl. Acad. Sci. USA **91**, 192-196

Guerini, D., Montell, C., and Klee, C.B. (1992). Molecular cloning and characterisation of the genes encoding the two subunits of *Drosophila melanogaster* calcineurin. J. Biol. Chem. **267**, 22542-22549.

Hall, J.C. (1994). The mating of a fly. Science **264**, 1702-1714.

Halpain, S., Girault, J.-A., Greengard, P. (1990). Activation of NMDA receptors induces dephosphorylation of DARPP-32 in rat striatal slices. *Nature* **343**, 369-372.

Hamilton, B.A., Palazzolo, M.J., Chang, J.H., Raghavan, K.V., Mayeda, C.A., Whitney, M.A., and Meyerowitz, E.M. (1991). Large scale screen for transposon insertions into cloned genes. *Proc. Natl. Acad. Sci. USA* **88**, 2731-2735.

Han, P.-L., Levin, R.L., Reed, R.R., and Davis, R.L. (1992). Preferential expression of the *Drosophila rutabaga* gene in mushroom bodies, neural centers for learning in insects. *Neuron* **6**, 619-627.

Han, P.-L., Meller, V., and Davis, R.L. (1996). The *Drosophila* Brain Revisited by Enhancer Detection. *J. Neurobiol.* **31**, 88-102.

Harding, M.W., Galat, A., Uehling, D.E., and Schreiber, S.L. (1989). A receptor for the immunosuppressant FK-506 is a *cis-trans* peptidyl-prolyl isomerase. *Nature* **341**, 761-763.

Hawkins, R.D., Abrams, T.W., Carew, T.J., and Kandel, E.R. (1983). A cellular mechanism of classical conditioning in *Aplysia*: activity-dependent amplification of presynaptic facilitation. *Science* **219**, 400-405.

Heisenberg, M. (1980). Development and Neurobiology of *Drosophila* (eds Siddiqi, Q. *et al*) 373-390 (Plenum, New York)

Heisenberg, M. (1989). Genetic approach to learning and memory (mnemogenetics) in *Drosophila melanogaster*. In 'Fundamentals of memory formation: neuronal plasticity and brain function.' (New York: Gustav Fischer Verlag) pp. 3-45.

Heitman, J., Movva, N.R., Hiestand, P.C., and Hall, M.N. (1991). FK 506-binding protein proline rotamase is a target for the immuno-suppressive agent FK506 in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **88**, 1948-1952.

Henrich, V.C., Sliter, T.J., Lubahn, D.B., Macintyre, A., and Gilbert, L.I. (1990). A steroid/thyroid hormone receptor superfamily member in *Drosophila melanogaster* that shares extensive sequence similarity with a mammalian homologue. *Nucleic Acids Res.* **18**, 4143-4148.

Henrich, V.C., Szekely, A.A., Kim, S.J., Brown, N.E., Antoniewski, C., Hayden, M.A., Lepesant, J.-A., and Gilbert, L.I. (1994). Expression and Function of the *ultraspiracle (usp)* Gene during development of *Drosophila melanogaster*. *Develop. Biol.* **165**, 38-52.

Hubbard, M.J., and Klee, C.B. (1989). Functional domain structure of calcineurin A: mapping by limited proteolysis. *Biochemistry* **28**, 1868-1874.

Immanuel, D., Zinszner, H., and Ron, D. (1995). Association of SARFH (Sarcoma-Associated RNA-Binding Fly Homolog) with regions of Chromatin Transcribed by RNA Polymerase II. *Mol. Cell. Biol.* **15**, 4562-4571.

Inoue, A., Takahashi, K., Kimura, M., Watanabe, T., and Morisawa, S. (1996). Molecular cloning of a RNA-binding protein, S1-1. *Nucleic Acids Res.* **24**, 2990-2997.

Isakov, N., Mally, M.I., Scholz, W., and Altman, A. (1987). Lymphocyte-T activation - the role of protein kinase-C and the bifurcating inositol phospholipid signal transduction pathway. *Immunol. Rev.* **95**, 89-111.

Ito, K., Awano, W., Suzuki, K., Hiromi, Y., and Yamamoto, D. (1997). The *Drosophila* mushroom body is a quadruple structure of clonal units each of which contains a virtually identical set of neurones and glial cells. *Development* **124**, 761-771.

Ito, K., and Hotta, Y. (1991). Proliferation pattern of postembryonic neuroblasts in the brain of *Drosophila melanogaster*. *Dev. Biol.* **149**, 134-148.

Jayaraman, T., Brilliantes, A.-M.B., Timerman, A.P., Erdjument-Bromage, H., Fleischer, S., Tempst, P., and Marks, A.R. (1992). FK506-binding protein associated with the calcium release channel (ryanodine receptor). *J. Biol. Chem.* **267**, 9474-9477.

Jones, C., and Tully, T. (1995). Anatomical dissection of associative and nonassociative learning in *Drosophila melanogaster*. *Abstr. Conf. Dros. Res.* **36**, 93B.

Kaiser, K., and Goodwin, S.F. (1990). "Site selected" transposon mutagenesis of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **87**, 1686-1690.

Kandel, E.R. and Schwartz, J.H. (1982). Molecular biology of an elementary form of learning: modulation of transmitter release by cAMP. *Science* **218**, 433-443.

Kater, S.B., Mattson, M., Cohan, C.S. and Connor, J. (1988). Calcium regulation of the neuronal growth cone. *Trends Neurosci.* **11**, 315-321.

Kauvar, L.M. (1982). Defective cyclic adenosine 3':5'-monophosphate phosphodiesterase in the *Drosophila* memory mutant *dunce*. *J. Neurosci.* **2**, 1347-1358.

Kim, Y.-T., and Wu, C.-F. (1996). Reduced growth cone motility in cultured neurons from *Drosophila* memory mutants with a defective cAMP cascade. *J. Neurosci.* **16**, 5593-5602.

Kishimoto, A., Kajikawa, N., Shiota, M., and Nishizuka, Y. (1983). Proteolytic activation of calcium-activated phospholipid-dependent protein-kinase by calcium-dependent neutral protease. *J. Biol. Chem.* **258**, 1156-1164.

Klee, C.B., Draetta, G., and Hubbard, M.J. (1987). Calcineurin. *Adv. Enzymol.* **61**, 149-200.

Klee, C.B., and Cohen, P. (1988). The calmodulin regulated protein phosphatase. In 'Calmodulin. Molecular aspects of cellular recognition.' (Amsterdam: Elsevier) pp. 225-248.

Koltin, Y., Faucette, L., Bergsma, D.J., Levy, M.A., Cafferkey, R., Koser, P.L., Johnson R.K., and Livi, G.P. (1991). Rapamycin sensitivity in *Saccharomyces cerevisiae* is mediated by a peptidyl-prolyl *cis-trans* isomerase related to human FK506-binding protein. *Mol. Cell. Biol.* **11**, 1718-1723.

Kretsinger, R.H. (1980). Crystallographic studies of calmodulin and homologs. *Ann. NY Acad. Sci.* **356**, 14-19.

Kunz, J., Henriquez, R., Schneider, U., Deuterreinhard, M., Mourra, N.R., and Hall, M.N. (1993). Target of rapamycin in yeast, TOR2, is an essential phosphatidylinositol kinase homolog required for G1 progression. *Cell* **73**, 585-596.

Lakey, A., Ferguson, C., Labeit, S., Reedy, M., Larkins, A., Butcher, G., Leonard, K., and Bullard, B. (1990). Identification and localization of high molecular weight proteins in insect flight and leg muscle. *EMBO J.* **9**, 3459-3467.

Lane, M.E., and Kalderon, D. (1993) Genetic investigation of cAMP-dependent protein kinase function in *Drosophila* development. *Genes Dev.* **7**, 1229-1243.

Lane, W.F., Galat, A., Harding, M.W., and Schreiber, S.L. (1991). Complete amino acid sequence of the FK506 and rapamycin binding protein, FKBP, isolated from calf thymus. *J. Protein Chem.* **10**, 151-160.

Lankford, K.L., and Letourneau, P. (1991). Roles of actin filaments and three second messenger systems in short term regulation of chick dorsal root ganglion neurite outgrowth. *Cell Motil. Cytoskel.* **8**, 193-209.

Lebeau, M.-C., Massol, N., Herrick, J., Faber, L.E., Renoir, J.-M., Radanyi, C., and Baulieu, E.-E. (1992). P59, an hsp 90-binding protein. Cloning and sequencing of its cDNA and preparation of a peptide-derived polyclonal antibody. *J. Biol. Chem.* **267**, 4218-4284.

Letourneau, P.C., Shattuck, T.A., and Ressler, A.H. (1987). 'Pull' and 'push' in neurite elongation: observations on the effects of different concentrations of cytochalasin B and taxol. *Cell Motil. Cytoskel.* **90**, 631-643.

Levi, A., Biocca, A., Cattaneo, A., and Calissano, P. (1988). The mode of action of nerve growth-factor in PC12 cells. *Mol. Neurobiol.* **2**, 201-226.

Levin, L.R., Han, P.-L., Hwang, P.M., Feinstein, P.G., Davis, R.L., and Reed, R.R. (1992). The *Drosophila* learning and memory gene *rutabaga* encodes a  $Ca^{2+}$ /calmodulin-responsive adenylyl cyclase. *Cell* **68**, 479-489.

Levin, J.M., Robson, B., and Garnier, J. (1986) An algorithm for secondary structure determination in proteins based on sequence similarity. *FEBS Lett.* **205**, 303-308

Levy, L.S., and Manning, J.E. (1981). Messenger RNA sequence complexity and homology in developmental stages of *Drosophila*. *Dev. Biol.* **85**, 141-149.

Lewis, A.K., and Bridgman, P.C. (1992). Nerve growth cone lamellipodia contain two populations of actin filaments that differ in organization and polarity. *J. Cell Biol.* **119**, 1219-1243.

Lindsley, D.L., and Zimm, G.G. (1992). *The Genome of Drosophila melanogaster*. San Diego, CA. Academic Press Inc.

Liu, J., Farmer Jr., J.D., Lane, W.S., Friedman, J., Weissman, I., and Schreiber, S.L. (1991). Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* **66**, 807-815.

Livingstone, M.S. (1985) Genetic dissection of *Drosophila* adenylyl cyclase. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 5992-5996.

Livingstone, M.S., Sziber, P.P., and Quinn, W.G. (1984). Loss of calcium calmodulin responsiveness in adenylyl-cyclase of *rutabaga*, a *Drosophila* learning mutant. *Cell* **37**, 205-215.

Lovering, R., Hanson, I.M., Borden, K.L.B., Martin, S., O'Reilly, N.J., Evan, G.I., Rahman, D., Pappin, D.J.C., Trowsdale, J., and Freemont, P.S. (1993). Identification and preliminary characterisation of a protein motif related to the zinc finger. *Proc Natl. Acad. Sci. USA* **90**, 2112-2116.

Lyons W.E., George, E.B., Dawson, T.M., Steiner, J.P., and Snyder, S.H. (1994). Immunosuppressant FK506 promotes neurite out-growth in cultures of PC12 cells and sensory ganglia. *Proc. Natl. Acad. Sci. USA* **91**, 3191-3195.

Lyons, W.E., Steiner, J.P., Snyder, S.H. and Dawson, T.M. (1995). Neuronal regeneration enhances the expression of the immunophilin FKBP-12. *J. Neurosci.* **15**, 2985-2994.

Malinow, R., Schulman, H., and Tsien, R.W. (1989). Inhibition of postsynaptic PKC or CaMKII blocks induction but not expression of LTP. *Science* **245**, 862-866.

Mattila, P.S., Ullman, K.S., Fiering, S., Emmel, E.A., McCutcheon, M., Crabtree, G.R., and Herzenberg, L.A. (1990). The actions of cyclosporin A and FK506 suggest a novel step in the activation of T lymphocytes. *EMBO J.* **9**, 4425-4431.

McMorrow, I., Bastos, R., Horton, H., and Burke, B. (1994). Sequence analysis of a cDNA encoding a human nuclear pore complex protein, hnup153. *Biochim Biophys Acta* **1217**, 219-223.

Mead, D.A., Skorupa, E.S., and Kemper, B. (1985). Single-stranded "Blue" T7 promoter plasmids: A versatile tandem promoter system for cloning and protein engineering. *Nucl. Acids Res.* **13**, 1103-1118.

Meissner, G. (1994). Ryanodine receptor Ca<sup>2+</sup> release channels and their regulation by endogenous effectors. *Annu. Rev. Physiol.* **56**, 485-508.

Michnick, S.W., Roen, M.K., Wandless, T.J., Karplus, M., and Schreiber, S.L. (1991). Solution structure of FKBP, a rotamase enzyme and receptor for FK506 and rapamycin. *Science* **252**, 836-839.

Mihalek, R., Jones, C., and Tully, T. (1995) Rescue of the lethal phenotype of *latheo*, a gene required for normal learning and memory in *Drosophila*. *Abstr. Conf. Dros. Res.* **36**, 98A.

Milligan, C.D. (1995). PhD Thesis, University of Glasgow.

Mitchell, P.J., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**, 371-378.

Moffat, K.G., Gould, J.H., Smith, H.K., and O'Kane, C.J. (1992). Inducible cell ablation in *Drosophila* by cold sensitive ricin A chain. *Development* **114**, 681-687.

Moriyoshi, K., Masu, M., Ishii, T., Shigemoto, R., Mizuno, N., and Nakanishi, S. (1991). Molecular cloning and characterisation of the rat NMDA receptor. *Nature* **354**, 31-37.

Morris, R.G.M., Anderson, E., Lynch, G.S., and Baundry, M. (1986). Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. *Nature* **319**, 774-776.

Mount, M.S., Burks, C., Hertz, G., Stormo, G.D., White, O., and Fields, C. (1992). Splicing signals in *Drosophila*: intron size information content, and consensus sequences. *Nucleic Acids Res.* **20**, 4255-4262.

Mulkey, R.M., Endo, S., Shenolikar, S., and Malenka, R.C. (1994). Involvement of a calcineurin/inhibitor-1 phosphatase cascade in hippocampal long-term depression. *Nature* **369**, 486-488.

Murray, N.E. (1983). Phage Lambda and Molecular Cloning. In *Lambda II*, 395-433. (Eds. Hendrix, R.W. *et al.*) Cold Spring Har-bour Laboratory Press.

Nagase, T., Seki, N., Tanaka, A., Ishikawa, K., and Nomura, N. (1995). Prediction of the coding sequences of unidentified human genes IV. The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by analysis of cDNA clones from human cell line KG-1. *DNA Res.* **2**, 167-174.

Nassel, D.R., and Elekes, K. (1992). Aminergic neurons in the brain of blowflies and *Drosophila*: dopamine- and tyrosine hydroxylase-immunoreactive neurons and their relationship with putative histaminergic neurons. *Cell Tissue Res.* **267**, 147-167.

Nelson, P.A., Lippke, J.A., Murcko, M.A., Rosborough, S.L., and Peattie, D.A. (1991). cDNA encoding murine FK506-binding protein (FKBP): nucleotide and deduced amino acid sequences. *Gene* **109**, 255-258.

Nichols, R.A., Suplick, G.R., and Brown, J.M. (1994). Calcineurin-mediated protein dephosphorylation in brain nerve-terminals regulates the release of glutamate. *J. Biol. Chem.* **269**, 23817-23823.

Nighorn, A., Healy, M.J., and Davis, R.L. (1991). The cyclic AMP phosphodiesterase encoded by the *Drosophila dunce* gene is concentrated in the mushroom body neuropil. *Neuron* **6**, 455-467.

O'Dell, K.M.C., Armstrong, J.D., Yang, M.Y., and Kaiser, K. (1995). Functional Dissection of the *Drosophila* Mushroom Bodies by Selective Feminization of Genetically Defined Subcompartments. *Neuron* **15**, 55-61.

O'Kane, C.J., and Gehring, W.J. (1987). Detection *in situ* of genomic regulatory elements in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **84**, 9123-9127.

Oliner, J.D., Kinzler, K.W., Meltzer, P.S., George, D.L., and Vogelstein, B. (1992). Amplification of a gene encoding a p53-associated protein in human sarcomas. *Nature* **358**, 80-83.

Ono, Y., Fujii, T., Igarashi, K., Kuno, T., Tanaka, C., Kikkawa, U., and Nishizuka, Y. (1989). Phorbol ester binding to protein kinase C requires a cysteine-rich zinc-finger-like sequence. *Proc. Natl. Acad. Sci. USA* **86**, 4868-4871.

Oro, A.E., McKeown, M., and Evans, R.M. (1990). Relationship between the product of the *Drosophila ultraspiracle* locus and the vertebrate retinoid X receptor. *Nature* **347**, 298-301.

Otey, C.A., Pavalco, F.M., and Burridge, K. (1990). An interaction between  $\alpha$ -actinin and the  $\beta_1$  integrin subunit *in vitro*. *J. Cell Biol.* **111**, 721-729.

Patel, N.H., and Goodman, C.S. (1992). Detection of *even-skipped* transcripts in *Drosophila* embryos with PCR/DIG-labelled probes. In 'Nonradioactive *in situ* hybridization application manual' Boehringer Mannheim Biochemica.

Pfeiffer-Linn, C., and Glantz, R.M. (1991). An arthropod NMDA receptor. *Synapse* **9**, 35-42.

Price, D.J., Grove, J.R., Calvo, V., Avruch, J., and Bierer, B.E. (1992). Rapamycin-induced inhibition of the 70-kilodalton S6 protein kinase. *Science* **257**, 973-977.

Prokop, A., and Technau, G.M. (1991). The origin of postembryonic neuroblasts in the ventral nerve cord of *Drosophila melanogaster*. *Development* **111**, 79-88.

Prokop, A., and Technau, G.M. (1994). Normal function of the mushroom body defect gene of *Drosophila* is required for the regulation of the number and proliferation of neuroblasts. *Dev. Biol.* **161**, 321-337.

Qui, Y., Chen, C-N, Malone, T., Richter, L., Beckendorf, S.K., and Davis, R.L. (1991). Characterisation of the memory gene *dunce* of *Drosophila melanogaster*. *J. Mol. Biol.* **222**, 553-565.

Qui, Y., and Davis, R.L. (1993). Genetic dissection of the learning/memory gene *dunce* of *Drosophila melanogaster*. *Genes Dev.* **7**, 1447-1458.

Quinn, W.G., Sziber, P.P., and Booker, R. (1979). The *Drosophila* memory mutant *amnesiac*. *Nature* **277**, 212-214.

Rees, C.T., and Spatz, H.-C. (1989). Habituation of the landing response of *Drosophila* wild-type and mutants defective in olfactory learning. *J. Neurogenet.* **5**, 105-118.

Rio, D.C., (1991). Regulation of *Drosophila* P-element transposition. *TIG* **7**, 282-287.

Robertson, H.M., Preston, C.R., Phillis, R.W., Johnson-Schlitz, D.M., Benz, W.K., and Engels, W.R. (1988). A stable source of P-element transposase in *Drosophila melanogaster*. *Genetics* **118**, 461-470.

Rosenmund, C., and Westbrook, G.L. (1993). Calcium-induced actin depolymerisation reduces NMDA channel activity. *Neuron* **10**, 805-814.

Rost, B., and Sander, C. (1994). Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins* **19**, 55-72.

Roulier, E.M., Fyrberg, C., and Fyrberg, E. (1992). Perturbations of *Drosophila*  $\alpha$ -actinin cause muscle paralysis, weakness, and atrophy but do not confer obvious nonmuscle phenotypes. *J. Cell Biol.* **116**, 911-922.

Sabatini, D.M., Erdjument-Bromage, H., Lui, M., Tempst, P., and Snyder, S.H. (1994). RAFT1 (Rapamycin and FKBP12 Target): a mammalian protein that binds to FKBP12 in rapamycin-dependent fashion and is homologous to yeast TORs. *Cell* **78**, 35-43.

Salz, H.K., Cline, T.W., and Schedl, P. (1987). Functional changes associated with structural alterations induced by mobilisation of a P element inserted in the *Sex-lethal* gene of *Drosophila*. *Genetics* **117**, 221-231.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular Cloning: A laboratory manual* (2nd ed.). Cold Spring Harbor Laboratory Press.

Sanchez-Garcia, I., and Rabbitts, T.H., (1994). The LIM domain: a new structural motif found in zinc-finger-like proteins. *TIG* **10**, 315-320.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463-5467.

Saurin, A.J., Borden, K.L.B., Boddy, M.N., and Freemont, P.S. (1996). Does this have a familiar RING? *TIBS* **21**, 208-214.

Schneuwly, S., Kuroiwa, A., Baumgartner, P., and Gehring, W.J. (1986). Structural organization and sequence of the homeotic gene *Antennapedia* of *Drosophila melanogaster*. *EMBO J.* **5**, 733-739.

Schreiber, S.L., and Crabtree, G.R. (1992). The mechanism of action of cyclosporin A and FK506. *Immunol. Today* **13**, 136-142.

Schwarz, T.L., Papazian, D.M., Carretto, R.C., Jan, Y.-N., and Jan, Y.L. (1990). Immunological characterisation of K<sup>+</sup> channel components from the *Shaker* locus and differential distribution of splicing variants in *Drosophila*. *Neuron* **2**, 119-127.

Seki, K., Chen, H.C., and Huang, K.P. (1995). Dephosphorylation of protein kinase C substrates, neurogranin, neuromodulin, and MARCKS, by calcineurin and protein phosphatases 1 and 2A. *Arch. Biochem. Biophys.* **316**, 673-679.

Sentry, J.W., and Kaiser, K. (1993). Harnessing the P-element transposon for studies of the *Drosophila* nervous system. *Current Topics in Mol. Gen.* **1**, 147-158.

Sentry, J.W., and Kaiser, K. (1994). Application of inverse PCR to site-selected mutagenesis of *Drosophila*. *Nucleic Acids Res.* **22**, 3429-3430.

Sentry, J.W., Goodwin, S.F., Milligan, C.D., Duncanson, A., Yang, M., and Kaiser, K. (1994). Reverse genetics of *Drosophila* brain structure and function. *Progress in Neurobiol.* **42**, 299-308.

Sewell, T.J., Lam, E., Martin, M.M., Leszyk, J., Weidner, J., Calaycay, J., Griffin, P., Williams, H., Hung, S., Cryan, J., Sigal, N.H., and Wiederrecht, G.J. (1994). Inhibition of Calcineurin by a novel FK-506-binding protein. *J. Biol. Chem.* **269**, 21094-21102.

Shotwell, S. (1983). Cyclic adenosine 3':5'-monophosphate phosphodiesterase and its role in learning in *Drosophila*. *J. Neurosci.* **3**, 739-747.

Siekerka, J.J., Hung, S.H.Y., Poe, M., Lin, S.C., and Sigal, N.H. (1989). A cytosolic binding protein for the immunosuppressant FK-506 has peptidyl-prolyl isomerase activity but is distinct from cyclophilin. *Nature* **341**, 755-757.

Skoulakis, E.M.C., Kalderon, D., and Davis, R.L. (1993). Preferential expression in mushroom bodies of the catalytic subunit of protein kinase A and its role in learning and memory. *Neuron* **11**, 197-208.

Smith, S.J. (1988). Neuronal cytom mechanics: the actin-based motility of growth cones. *Science* **242**, 708-715.

Sobue, K. and Kanda, K. (1989).  $\alpha$ -Actinins, caldesmons (brain spectrin or fodrin), and actin participate in adhesion and movement of growth cones. *Neuron* **3**, 311-319.

Standaert, R.F., Galat, A., Verdine, G.L., and Schreiber, S.L. (1990). Molecular cloning and overexpression of the human FK506-binding protein, FKBP. *Nature* **346**, 671-674.

Steiner, J.P., Dawson, T.M., Fotuhi, M., Glatt, C.E., Snowman, A.M., Cohen, N., and Snyder, S.H. (1992). High brain densities of the immunophilin FKBP colocalized with calcineurin. *Nature* **358**, 584-587.

Stewart, A.A., Ingebritsen, T.S., Manalan, A., Klee, C.B., and Cohen, P. (1982). Discovery of a  $\text{Ca}^{2+}$ - and calmodulin-dependent protein phosphatase. *FEBS Lett.* **137**, 80-84.

Stocker, R.F., Lienhard, M.C., Borst, A., and Fischbach, K.F. (1990). Neuronal architecture of the antennal lobe in *Drosophila melanogaster*. *Cell Tissue Res.* **262**, 9-34.

Stolow, D.T., and Haynes, S.R. (1995). *Cabeza*, a *Drosophila* gene encoding a novel RNA binding protein, shares homology with EWS and TLS, two genes involved in human sarcoma formation. *Nucleic Acids Res.* **23**, 835-843.

Strausfeld, N.J. (1976). Atlas of an insect brain. Springer-Verlag, Berlin.

Strauss, R., and Heisenberg, M. (1993). A higher control center of locomotor behavior in the *Drosophila* brain. *J. Neurosci.* **13**, 1852-1861.

Sukegawa, J., and Blobel, G. (1993). A nuclear pore complex protein that contains zinc finger motifs, binds DNA, and faces the nucleoplasm. *Cell* **72**, 29-38.

Taira, M., Jamrich, M., Good, P.J., and Dawid, I.B. (1992). The LIM domain-containing homeo-box gene *XLIM-1* is expressed specifically in the organizer region of *Xenopus* gastrula embryos. *Genes Dev.* **6**, 2165-2176.

Technau, G., and Heisenberg, M. (1982). Neural reorganization during metamorphosis of the *corpora pedunculata* in *Drosophila melanogaster*. *Nature* **295**, 405-407.

Tempel, B.L., Bonini, N., Dawson, D.R., and Quinn, W.G. (1983). Reward learning in normal and mutant *Drosophila*. *Proc. Natl. Acad. Sci. USA* **80**, 1482-1486.

Tempel, B.L., Livingstone, M.S., and Quinn, W.G. (1984). Mutations in the dopa decarboxylase gene affect learning in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **81**, 3577-3581.

Theopold, U., Zotto, L.D., and Hultmark, D., (1995). FKBP39, a *Drosophila* member of a family of proteins that bind the immunosuppressive drug FK506. *Gene* **156**, 247-251.

Timerman, A.P., Ogunbumni, E., Freund, E., Wiedderrecht, G., Marks, A.R., and Fleischer, S. (1993). The calcium release channel of sarcoplasmic reticulum is modulated by FK-506-binding protein. *J. Biol. Chem.* **31**, 22992-22999.

Timerman, A.P., Wiedderrecht, G., Marcy, A., and Fleischer, S. (1995). Characterization of an exchange reaction between soluble FKBP12 and the FKBP Ryanodine receptor complex. Modulation by FKBP mutants deficient in peptidyl-prolyl isomerase activity. *J. Biol. Chem.* **270**, 2451-2459.

Tocci, M.J., Matkovich, D.A., Collier, K.A., Kwok, P., Dumont, F., Lin, S., Degudicibus, S., Siekierka, J.J., Chin, J., and Hutchinson, N.I. (1989). The immunosuppressant FK506 selectively inhibits expression of early T cell activation genes. *J. Immunol.* **143**, 718-726.

Tong, G., Shepherd, D., and Jahr, C. (1995). Synaptic desensitization of NMDA receptors by calcineurin. *Science* **267**, 1510-1512.

Török, T., Tick, G., Alvarado, M., and Kiss, I. (1993). *P-lacW* insertional mutagenesis on the 2nd chromosome of *Drosophila melanogaster*: Isolation of lethals with different overgrowth phenotypes. *Genetics* **135**, 71-80.

Tropschug, M., Wachter, E., Mayer, S., Schonbrunner, E.R., and Schmid F.X. (1990). Isolation and sequence of an FK506-binding protein from *N. crassa* which catalyses protein folding. *Nature* **343**, 674-677.

Truman, J.W. (1990). Metamorphosis of the central nervous system of *Drosophila*. *J. Neurobiol.* **21**, 1072-1084.

Tsubota, S., and Schedl, P. (1986). Hybrid dysgenesis-induced revertants of insertions at the 5' end of the *rudimentary* gene in *Drosophila melanogaster*: transposon-induced control mutations. *Genetics* **114**, 165-182.

Tully, T., and Quinn, W.G. (1985). Classical conditioning and retention in normal and mutant *Drosophila melanogaster*. *J. Comp. Physiol.* **157**, 263-277.

Tully, T., Cambiazo, V., and Kruse, L. (1994). Memory through metamorphosis in normal and mutant *Drosophila*. *J. Neurosci.* **14**, 68-74.

Tully, T., Preat, T., Boynton, S.C., and Del Vecchio, M. (1994) Genetic dissection of consolidated memory in *Drosophila*. *Cell*, **79**, 35-47.

Tully, T., Boynton, S., Brandes, C., Dura, J.-M., Mihalek, R., Preat, T., and Vilella, A. (1990). Genetic dissection of memory formation in *Drosophila melanogaster*. *Cold spring Harbor Symp. Quant. Biol.* **55**, 203-211.

Ulloa, L., Dombradi, V., Diaznido, J., Szucs, K., Gergely, P., Friedrich, P., and Avila, J. (1993). Dephosphorylation of distinct sites of microtubule-associated protein MAP1B by protein phosphatases 1, 2A and 2B. *FEBS Lett.* **330**, 85-89.

Ultsch, A., Schuster, C.M., Laube, B., Betz, H., and Schmitt, B. (1993). Glutamate receptors of *Drosophila melanogaster*. Primary structure of a putative NMDA receptor protein expressed in the head of the adult fly. *FEBS Lett.* **324**, 171-177.

van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., and Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* **229**, 105-124.

Wang, J.H., and Stelzer, A. (1994). Inhibition of phosphatase 2B prevents expression of hippocampal long-term potentiation. *Neuroreport* **5**, 2377-2380.

Wang, T., Donahoe, P.K., and Zervos, A.S. (1994). Specific interaction of type I receptors of the TGF- $\beta$  family with the immunophilin FKBP-12. *Science* **265**, 674-676.

Wang, T., Li B.-Y., Danielson, P.D., Shah, P.C., Rockwell, S., Lechleider, R.J., Martin, J., Manganaro, T., and Donahoe, P.K. (1996). The immunophilin FKBP12 functions as a common inhibitor of the TGF- $\beta$  family type I receptors. *Cell*, **86** 435-444.

Wang, X., Lee, G., Liebhaber, S.A., and Cooke, N.E. (1992). Human cysteine-rich protein - a member of the LIM double-finger family displaying coordinate serum induction with c-myc. *J. Biol. Chem.* **267**, 9176-9184.

Way, J.C. and Chalfie, M. (1988). *Mec-3*, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in *C.elegans*. *Cell* **54**, 5-16

Whittaker, P.A., Campbell, A.J.B., Southern, E.M., and Murray, N.E. (1988). Enhanced recovery and restriction mapping of DNA fragments cloned in a new lambda vector. *Nucl. Acids Res.* **16**, 6725-6736.

Wilken, N., Senecal, J.L., Scheer, U., and Dabauvalle, M.C. (1995). Localization of the Ran-GTP binding protein RanBP2 at the cytoplasmic side of the nuclear pore complex. *Eur. J. Cell Biol.* **68**, 211-219.

Wright, N.J.D. and Zhong, Y. (1995). Characterisation of K<sup>+</sup> currents and the cAMP dependent modulation in cultured *Drosophila* mushroom body neurons identified by *lacZ* expression. *J. Neurosci.* **15**, 1025-1034.

Wyszynski, M., Lin, J., Rao, A., Nigh, E., Beggs, A.H., Craig, A.M., and Sheng, M. (1997). Competitive binding of  $\alpha$ -actinin and calmodulin to the NMDA receptor. *Nature* **385**, 439-442.

Yang, M.Y. (1995). PhD Thesis, University of Glasgow.

Yang, M.Y., Armstrong, J.D., Vilinsky, I., Strausfeld, N.J., and Kaiser, K. (1995). Subdivision of the *Drosophila* Mushroom Bodies by Enhancer-Trap Expression Patterns. *Neuron* **15**, 45-54.

Yao, T.-P., Forman, B.M., Jiang, Z., Cherbas, L., Chen, J.-D., McKeown, P.C., and Evans, R.M. (1993). Functional ecdysone receptor is the product of *EcR* and *Ultraspiracle* genes. *Nature* **366**, 476-479.

Yokoyama, N., Hayashi, N., Seki, T., Pante, N., Ohba, T., Nishii, K., Kuma, K., Hayashida, T., Miyata, T., Aebi, U., Fukui, M., and Nishimoto, T. (1995). A giant nucleopore protein that binds Ran/TC4. *Nature* **376**, 184-188.

